

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 00:56:52 ; Search time 405 Seconds

(without alignments)
6856.651 Million cell updates/sec

Title: US-10-725-829-1

Perfect score: 529
Sequence: 1 cataatggctgttcacaag.....ctgtaatggactttattg 529

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	529	100.0	529	12	ADP02854
2	288.4	54.5	652	12	ADP02856
3	259.2	49.0	684	10	ADP05323
4	258	48.8	1241	2	AAQ02053
5	254.2	48.1	584	3	Aaz49862
6	178.8	33.8	1360	2	AAQ68729
7	149	28.2	1104	2	AAQ68728
8	149	28.2	1104	6	ADI56863
9	35.8	6.8	2000	8	ADA71938
10	35.6	6.7	48727	4	AAK67375
11	35.6	6.7	108359	9	ADA13316
12	35.2	6.7	6134	4	AAS46563
13	35.2	6.7	6650	6	ABL32638
14	35.2	6.7	10433	6	ABL232379
15	35	6.6	9964	6	ABL32098
16	34.8	6.6	2591	4	ABL08938
17	34.8	6.6	6866	6	ABL32666
18	34.8	6.6	6866	6	ABL49319
19	34.6	6.5	346	6	ABN94362
20	34.4	6.5	2905	12	ADJ75932
21	34.4	6.5	63155	10	ADC85996

22	34.4	6.5	110000	6	ABA90521.19	Continuation (20 o
23	34.4	6.5	193303	12	ADP13122	Adf13122 Hypermeth
24	34.4	6.5	193303	12	ADP13115	Adf13115 Hypermeth
25	34.4	6.5	193303	12	ADI37268	Adi37268 Hypermeth
26	34.4	6.5	193303	12	ADI37261	Adi37261 Hypermeth
27	34.2	6.5	1262	2	AAX13696	Aax13696 Enterococ
28	34.2	6.5	1262	6	ABS99491	Abse9491 Enterococ
29	34	6.4	2864	4	ABL04388	AbL04388 Drosophil
30	34	6.4	32249	4	ABL04789	AbL04789 Human rep
31	34	6.4	32249	4	ABL97684	AbL97684 Human tes
32	34	6.4	75899	6	ABK85261	Abk85261 Human gen
33	34	6.4	75899	12	ADI13990	Adi13990 Human pro
34	33.8	6.4	5473	4	AAI58381	Aai58381 Human pol
35	33.8	6.4	5473	5	ADQ98590	Adq98590 DNA encod
36	33.8	6.4	5473	9	ADB48350	AdB48350 Novel hum
37	33.6	6.4	7545	8	AAI50282	Aai50282 Human nuc
38	33.6	6.4	672	10	ADC76022	Adc76022 DNA homol
39	33.6	6.4	707	10	ADK55393	Adk55393 Plant DNA
40	33.6	6.4	4043	4	AAH53990	Aah53990 S. epider
41	33.6	6.4	13578	4	AAK81662	Aak81662 Human imm
42	33.6	6.4	48203	4	AAK82628	Aak82628 Human imm
43	33.6	6.4	48203	4	AAK81663	Aak81663 Human imm
44	33.6	6.4	48203	4	AAK70161	Aak70161 Human imm
45	33.6	6.4	48204	4	AAK81666	Aak81666 Human imm

ALIGNMENTS

RESULT 1
ADP02854
ID ADP02854 standard; DNA; 529 BP.
XX
AC ADP02854;
XX
DT 09-SEP-2004 (first entry)
XX
DE Solanum americanum proteinase inhibitor II gene SapIN2a.
XX
KW ds; gene; proteinase inhibitor II; transformed plant; resistance; insect;
KW pest; pathogen; programmed cell death; senescence.
XX
OS Solanum americanum.
XX
FH Key Location/Qualifiers
CDS 5..451
FT /*tag= a
FT /product= "SapIN2a protein"

WO2004050873-A1.
XX
PD 17-JUN-2004.
XX
PF 01-DEC-2003; 2003WO-CN001020.
XX
PR 29-NOV-2002; 2002US-0429992P.
XX
PA (UYHK-) UNIV HONG KONG.
XX
PI Chye M, Xu Z, Sin S;
XX
DR WPI; 2004-450731/42.
XX
PT P-PSDB; ADP02855.
XX
PT New proteinase inhibitor, SapIN2a or SapIN2b, useful in producing
PT transformed plants having enhanced resistance to insects, pest or
PT pathogens and in which programmed cell death or senescence is inhibited.
XX
PS Claim 1; SEQ ID NO 1; 90pp; English.
XX
CC The invention relates to an isolated proteinase inhibitor II nucleic acid
CC molecule (I). The proteinase inhibitor II nucleic acid molecule and the
CC encoded polypeptide, methods are useful in producing transformed plants

QY 131 CCAGTTTCAGAGGAAGTCCCAAAACCTATATGCACCAATTTGTTGCTCAGGCTATAAG 190
DB |||||
558 CCAGTTTCAGAGGAAGTCCGGAATCCGATATGCACCAACTGTTGTGCAGGTTATAA 617
QY 191 GGTTCGAACTTACAGTCTAAAGGAGATTGATTTGGAAGGAGATCTGACCCCTAGA 250
DB |||||
618 GGTTCGAACTTATATAGTGCATAATGGGGCTTTCATTTGTGAAGGACAATCTGACCCAAA 677
QY 251 AACCCAAAGATTGACCTTCGAAATGTGATACACAGATTCTTATTCAAAATGTCTCGT 310
DB |||||
678 AAACCAAAAGCATGCCCTTAATTTGGATGCCATATTCCTACTCAAAAGTGTCCCGT 737
QY 311 TCAGAAGGAAGATGATAATTAACCCACCTGGATGCACCACTTTGTTGCACGGGCTATCAG 370
DB |||||
738 TCAGAAGGAATCGCTAATTTATCCACCGGATGTACCAACATGCTGCACAGGGTACAAG 797
QY 371 GGTTCGCTACTATTTTCGATCAAGATGGTGATTTGCTGTGAAGGAGAGTCTGAAACC 430
DB |||||
798 GGTTCGCTACTATTTTCGCTAAAAATGGCAAGTTTGTATGTGAAGGAGAGTGTAGGCC 857
QY 431 AAGCACTGCTTA 444
DB |||||
858 AAGCAAAATATGTA 871

RESULT 5
AAZ49862
ID AAZ49862 standard; DNA; 584 BP.
XX AC AAZ49862;
XX AC AAZ49862;
DT 25-APR-2000 (first entry)
XX DE Potato proteinase inhibitor-II DNA.
XX KW Potato proteinase inhibitor-II; PPI-II; streptavidin; worm; insect;
KW plant-noxious protein; pest resistance; moth; insect; weevil; grub;
KW beetle; fly; thrip; locust; cricket; borer; mite; looper; insecticidal;
KW ss.
XX OS Solanum tuberosum.
XX FH Key Location/Qualifiers
FT CDS 1..584
FT FT /*tag= a
FT FT /product= "Potato proteinase inhibitor-II"
FT FT /notes= "coding region contains one intron"
FT FT sig_peptide 1..212
FT FT /*tag= b
FT FT /note= "signal peptide contains one intron"
FT FT exon 1..52
FT FT /*tag= d
FT FT /number= 1
FT FT intron 53..171
FT FT /*tag= d
FT FT /number= 1
FT FT exon 172..584
FT FT /*tag= d
FT FT /number= 2
FT FT mat_peptide 213..581
FT FT /*tag= c
XX WO200004049-A1.
XX 27-JAN-2000.
XX 15-JUL-1999; 99WO-NZ000110.
XX 15-JUL-1998; 98NZ-00331002.
XX (HORT-) HORTICULTURE & FOOD RES INST NEW ZEALAND.
XX PA
XX KW

PI Christeller JT, Sutherland PW, Murray C, Markwick NP, Philip BA;
PI Malone LA, Burgess BPJ;
XX WPI; 2000-171244/15.
XX
PT New chimeric polypeptide and composition comprising the polypeptide
PT useful for conferring pest resistance on plants.
XX
PS Example 3; Fig 4; 11pp; English.
XX
CC The present sequence encodes potato proteinase inhibitor-II (PPI-II).
CC This is used in the preparation of a binary vector designed to express a
CC chimeric polypeptide comprising streptavidin mature peptide, a plant -
CC noxious protein, fused to the PPI-II signal peptide. The binary vector is
CC targeted to the vacuole by PPI-II signal sequence. Transformation of
CC plant genome with the vector can produce pest resistance in plants, plant
CC derived products and stored harvest material. Pests that can be
CC controlled include, cotton bollworm, tropical army-worm, European corn -
CC borer or red mite, tobacco horn worm, looper, rice stem borer, porina,
CC cutworms, diamondback moth, potato tuber moth, codling moth, Indian meal
CC moth, gypsy moth, argentine stem weevil, clover root weevil, grass -
CC grubs, corn rootworm, rice and wheat weevils, mealworms, flour beetles,
CC black field cricket, locusts, sawflies, Western flower thrips, Hessian
CC flies or two-spotted mite
XX
SQ Sequence 584 BP; 184 A; 94 C; 116 G; 190 T; 0 U; 0 Other;
Query Match 48.1%; Score 254.2; DB 3; Length 584;
Best Local Similarity 79.4%; Pred.No.5.7e-66;
Matches 301; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 66 TTGCGAAACATGTTGATGCGCAAGGCTTGTAAGAGATGTGTCATTTTAGTATGGCA 125
DB |||||
193 TGGTGGAAACATGTTGATGCGCAAGATCTGTACTAAAGATGTGTAATCTTGGGTTGGGA 252
QY 126 TATGCCCAAGTTCAGAGGAAGTCCCAAAACCTATATGCACCAATTTGTTGCTCAGGCT 185
DB |||||
253 TATGCCCAAGTTCAGAGGAAGTCCGAAAATCCCATATGCATCAATTTGTTGCTCAGGCT 312
QY 186 ATAAGGGTTGCAACTATTACAGTGTCTAAAGAGATTTGATTTTGTGAAGGAGAACTTGACC 245
DB |||||
313 ATAAGGGTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 372
QY 246 CTAGAAACCAAAAGATTTGATCTTCGAATGTGATACACAGATTTGCTATTTCAAAATGTC 305
DB |||||
373 TAAATAACCAAAAGCTTCCCTCTAAATTTGTGATACAAATATTGCTATTCAAGATGCC 432
QY 306 CTCGTTTCAGAGGAAGATGATAATTAACCCCACTGGATGCACCACTTGTTCACCGGCT 365
DB |||||
433 CCCATTTCAGAGGAAGAAATCGCTAAATTTATCCCAACCGGATGTACCATGTTGCACAGGGT 492
QY 366 ATCAGGGTTGCTACTATTTCGATCAAGATGTGATTTTGTCTGTGAAGGAGAGAGTCCCTG 425
DB |||||
493 ACAGGGTTGCTACTATTTCGGTAAATAATGGCAAGTTTGTATCGGAAGAGAGAGATGATG 552
QY 426 AACCCAAAGCACTGCTTA 444
DB |||||
553 AACCCAAAGCAAAATATGTA 571
RESULT 6
AAQ68729
ID AAQ68729 standard; DNA; 1360 BP.
XX AC AAQ68729;
XX AC AAQ68729;
DT 25-MAR-2003 (revised)
DT 02-MAR-1995 (first entry)
XX
DE Full length sequence of PI precursor.
XX Type II serine proteinase inhibitor precursor; PI; tobacco;
KW transgenic plant; anti-pathogen; anti-predator; ss.

```
XX OS Nicotiana alata.
XX
XX FH Key Location/Qualifiers
XX FT misc_signal 1..97
XX FT /tag= b
XX FT /label= signal sequence
XX FT CDS 97..1200
XX FT /tag= a
XX
XX PN WO9413810-A1.
XX PN 23-JUN-1994.
XX
XX PF 16-DEC-1993; 93WO-AU000659.
XX
XX PR 16-DEC-1992; 92AU-00006399.
XX
XX PA (UYME ) UNIV MELBOURNE.
XX
XX PI Anderson MA, Atkinson AH, Heath RL, Clarke AE;
XX
XX DR WPI; 1994-217886/26.
XX DR P-PSDB; AAR54135.
XX
XX PT Nicotiana alata type II serine protease inhibitor precursor and DNA -
XX PT useful in prodn of anti-pathogen or anti-predator constructs for plants.
XX
XX PS Disclosure; Page 45-47; 83pp; English.
XX
XX CC A cDNA library, prepd. from mRNA from the stigmas and styles of mature
XX CC flowers of N. alata was screened for clones of highly expressed genes
XX CC which were not associated with self-incompatibility genotype. Clones
XX CC encoding a protein with some identity to the type II proteinase
XX CC inhibitors from potato and tomato were selected. The largest clone, NA-PI
XX CC -2, is given in AAQ68729. The predicted AA sequence in AAR54135. (Updated
XX CC on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 1360 BP; 416 A; 219 C; 345 G; 380 T; 0 U; 0 Other;
XX
XX Query Match 33.8%; Score 178.8; DB 2; Length 1360;
XX Best Local Similarity 67.3%; Pred. No. 3.9e-43;
XX Matches 303; Conservative 0; Mismatches 132; Indels 15; Gaps 3;
XX
QY 5 ATGGCTGTTTCAAAAGTTAGTTCCTTGGTGGCTTACTTGTTCCTGGATGGAGTTTCTA 64
DB 10 ATGGCTGTTTCAAAAGTTAGTTCCTTGGTGGCTTACTTGTTCCTGGATGGAGTTTCT 69
QY 65 CT-----TGCAGAACATGTTGATGCCAGGCTTGTACTAGAAATGTGGTCAAT---TT 115
DB 70 GTAAGCAATGTGGAAACATGCAGATGCCAAGGCTTGTACCTTAAACTGTGATCCAGAAT 129
QY 116 AGCTATGGCATATGCCACGTTTCAGAGGAAGTCCCAAAACCTATATGCACCAATTTGT 175
DB 130 GCCTATGGAGTTTGCCGCGTTTCAGAGAAAGAAAGAAATGATCGGATATGCACCACTGT 189
QY 176 TGCTCAGGCTATAGGGTTGCAACTATTACAGTGTCTAAAGGAGATTGATTGTGGAAGGA 235
DB 190 TGGCAGGACGCAAGGGTTGTAAGTACTTCAGTGATGATGGAACCTTTTGTGTGGAAGGA 249
QY 236 GAATCTGACCTAGAACCCAAAGATTGTACCTTGAATGTGATACACAGATTGCTTAT 295
DB 250 GAGTCTGATCTAGAAATCCAAAGGCTTGTACCTTAAACTGTGATCCAAAGAAATTCGCTAT 309
QY 296 TCAAAATGTCCTGTTTCAGAGGAAGATGATAATTAACCCACTGGATGCACCACTTGT 355
DB 310 GGAGTTTGCCCGCTTCAGAGAAAGAAAG------AATGATCGGATATGCACCACTGT 363
QY 356 TGCACGGGCTATCAGGGTTGCTACTATTTTCGATCAAGAGTGGTGAATTTTGTGTGGAAGGA 415
DB 364 TGCAGGACGCAAGGGTTGTAAGTACTTCAGTGATGATGGAACCTTTTGTGTGGAAGGA 423
QY 416 GAGAGTCTGAAACCCAAAGACCACTGCTTAT 445
```

```
|||||
424 GAGTCTGATCCTAGAAATCCAAAGGCTTGT 453

RESULT 7
AAQ68728
ID AAQ68728 standard; DNA; 1104 BP.
XX
XX AC AAQ68728;
XX
XX DT 25-MAR-2003 (revised)
XX DT 02-MAR-1995 (first entry)
XX
XX DE Nucleotide coding region of N-alata PI precursor.
XX
XX KW Type II serine proteinase inhibitor precursor; PI; tobacco;
XX KW transgenic plant; anti-pathogen; anti-predator; ss.
XX
XX OS Nicotiana alata.
XX
XX FH Key Location/Qualifiers
XX FT CDS 1..1104
XX FT /*tag= a
XX
XX PN WO9413810-A1.
XX
XX PD 23-JUN-1994.
XX
XX PF 16-DEC-1993; 93WO-AU000659.
XX
XX PR 16-DEC-1992; 92AU-00006399.
XX
XX PA (UYME ) UNIV MELBOURNE.
XX
XX PI Anderson MA, Atkinson AH, Heath RL, Clarke AE;
XX
XX DR WPI; 1994-217886/26.
XX DR P-PSDB; AAR54135.
XX
XX PT Nicotiana alata type II serine protease inhibitor precursor and DNA -
XX PT useful in prodn of anti-pathogen or anti-predator constructs for plants.
XX
XX PS Claim 5; Page 44-45; 83pp; English.
XX
XX CC A cDNA library, prepd. from mRNA from the stigmas and styles of mature
XX CC flowers of N. alata was screened for clones of highly expressed genes
XX CC which were not associated with self-incompatibility genotype. Clones
XX CC encoding a protein with some identity to the type II proteinase
XX CC inhibitors from potato and tomato were selected. The largest clone, NA-PI
XX CC -2, is given in AAQ68729. The predicted AA sequence in AAR54135. AAQ68728
XX CC is the coding region of AAQ68729. A nucleic acid isolate having at least
XX CC 55% similarity to AAQ68728 is claimed. (Updated on 25-MAR-2003 to correct
XX CC PN field.)
XX
XX SQ Sequence 1104 BP; 337 A; 176 C; 295 G; 296 T; 0 U; 0 Other;
XX
XX Query Match 28.2%; Score 149; DB 2; Length 1104;
XX Best Local Similarity 67.2%; Pred. No. 3.3e-34;
XX Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;
XX
QY 86 AAGGCTTGCTAGAGAAATGTGTC---ATTTAGCTATGGCATATGCCACGCTTCAGAA 142
DB 1 AAGGCTTGCTAGAGAAATGTGTC---ATTTAGCTATGGCATATGCCACGCTTCAGAA 60
QY 143 GGAAGTCCCAAAACCTATATGCACCAATTTGTCTCAGGCTATAGGGTTGCAACTAT 202
DB 61 GAAAGAGAAATGATCGGATATGCACCACTGTTCGCGAGGACGAGGGTTGTAAGTAC 120
QY 203 TACAGTGTCTAAAGGAGATTGTTGTGTAAGGAGAAATCTGACCTAGAAACCCAAAGAT 262
DB 121 TTCAGTGTATGATGGAACCTTTTGTGTAAGGAGAGTCTGATCTAGAAATCCAAAGGCT 180
QY 263 TGTACCTTCGAATGTGATACACAGATTGCTTATTCAAAATGTCCTCGTTCAGAGGAAG 322
```

```

||||| 181 TGTACCTTAAACTGTGATCAAGAAATTGCCATATGGAGTTTCCCGCGTTTCAGAGAAAG 240
||||| 323 ATGATAATTAAACCCACTGGATGCACCACTTGTTCACCGGCTATCAGGGTTCGTACTAT 382
||||| 241 AAG-----AATGATCGGATATGACCACTGTTCGCGCAGGACGAGGGTTGTAAGTAC 294
||||| 383 TTCGATCAAGATGGTGAATTTGTCTGTGAAGGAGAGAGTCTCTGAACCCCAAGACCACTGCT 442
||||| 295 TTCAGTCATGATGGAACCTTTGTTGTGAAGGAGAGTCTGATCCTAGAAATCCAAAGGCT 354
443 TAT 445
355 TGT 357

RESULT 8
ADI56863
ID ADI56863 standard; DNA; 1104 BP.
XX AC ADI56863;
XX DE 22-APR-2004 (first entry)
XX DE Plant defensin NAPI mature domain cDNA.
XX ds; gene; insecticide; antifungal; virucide; antibacterial;
XX flower development; plant pest resistance; floral defensin precursor;
XX transgenic plan; insect resistance.
XX Unidentified.
XX WO200263011-A1.
XX PD 15-AUG-2002.
XX PF 08-FEB-2002; 2002WO-AU000123.
XX PR 08-FEB-2001; 2001US-0267271P.
XX PA (HEXI-) HEXIMA LTD.
XX Anderson MA, Lay FT, Heath RL;
XX WPI; 2002-657538/70.
XX P-PSDB; ADI56864.
XX New isolated plant floral defensin-like polypeptides and polynucleotides,
XX for generating transgenic plants having resistance or at least reduced
XX sensitivity to plant pests including insects, microorganisms, fungi
XX and/or viruses.
XX Disclosure; SEQ ID NO 56; 164pp; English.
XX The invention relates to an isolated polypeptide comprising, in its
XX precursor form, an N-terminal signal domain, a mature domain and an
XX acidic C-terminal domain, where the polypeptide is produced during flower
XX development and its mature domain has activity against one or more plant
XX pests. Plant floral defensin-like polypeptides and polynucleotides are
XX useful in generating transgenic plants having resistance or at least
XX reduced sensitivity to plant pests including insects, microorganisms,
XX fungi and/or viruses. They are also useful in generating recombinant
XX defensin-like molecules for use in the topical application of
XX compositions to prevent or retard pest-infestation of plants. The floral-
XX and seed-derived defensins are useful in the generation of insect
XX resistance in plants. This sequence represents the cDNA encoding the
XX mature domain from a plant defensin protein.
XX Sequence 1104 BP; 337 A; 176 C; 295 G; 296 T; 0 U; 0 Other;
XX Query Match 28.2%; Score 149; DB 6; Length 1104;
XX Best Local Similarity 67.2%; Pred. No. 3.3e-34;
XX Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

```

```

QY 86 AAGGCTTGTACTAGAGAATGTGTC---ATTTTAGCTATGGCATATGCCACCGTTTCAGAA 142
||||| 1 AAGGCTTGTACTTAAACTGTGATCAAGAAATTGCCATATGGAGTTTCCCGCGTTTCAGAA 60
||||| 143 GGAAGTCCCAAAAACCTATATATGCACCAATTTGTTCAGGCTATTAAGGGTTGCAACTAT 202
||||| 61 GAAAAGAGAATGATCGGATATGCACCACTGTTCGCGCAGGACGAGGGTTGTAAGTAC 120
QY 203 TACAGTGTCTAAAGGAGATTTGATTTCTGAAGGAGAAATCTGACCCCTAGAAACCCCAAAAGAT 262
||||| 121 TTCAAGTGTGATGAGAACTTTTGTGTAAGGAGAGTCTGATCCTAGAAATCCAAAGGCT 180
QY 263 TGTACCTTCCAATGTGATACACAGATTGCTTATTTCAAAATGTCTCTGTTTCAGAGGAAAG 322
||||| 181 TGTACCTTAAACTGTGATCAAGAAATTGCTATGGAGTTTCCCGCGTTTCAGAGGAAAG 240
QY 323 ATGATAATTAAACCCACTGGATGCACCACTTGTTCACCGGCTATCAGGGTTCGTACTAT 382
||||| 241 AAG-----AATGATCGGATATGCACCACTGTTCGCGCAGGACGAGGGTTGTAAGTAC 294
QY 383 TTCGATCAAGATGGTGAATTTGTCTGTGAAGGAGAGTCTCTGAACCCCAAGACCACTGCT 442
||||| 295 TTCAAGTGTGATGGAACCTTTTGTGTAAGGAGAGTCTGATCCTAGAAATCCAAAGGCT 354
443 TAT 445
355 TGT 357

RESULT 9
ADA71938/c
ID ADA71938 standard; DNA; 2000 BP.
XX AC ADA71938;
XX DT 20-NOV-2003 (first entry)
XX DE Rice gene, SEQ ID 5263.
XX DE Plant; bacterial infection; fungal infection; viral infection; rice;
XX KW gene; ds.
XX OS Oryza sativa.
XX PN WO2003000898-A1.
XX PD 03-JAN-2003.
XX PF 22-JUN-2001; 2001WO-IB001105.
XX PR 22-JUN-2001; 2001WO-IB001105.
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
XX Identifying at least one gene involved in plant resistance or response to
XX pathogenic infection for conferring resistance or tolerance to a plant to
XX bacterial, fungal or viral infection by determining or detecting plant
XX gene expression.
XX Claim 27; SEQ ID NO 5263; 899pp; English.
XX The present invention relates to a method (M1) for identifying genes
XX involved in plant resistance or response to pathogenic infection. M1
XX comprises identifying a gene whose expression is significantly altered in
XX the incompatible interaction of plant gene expression relative to
XX expression of the gene in an uninfected plant, in a mutant plant that
XX does not express a gene associated with response to pathogenic infection,

```


PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246479P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0254097P.
PR 03-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
DR
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
PT
XX
XX Disclosure; SEQ ID NO 22187; 307lpp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703

CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 48727 BP; 16775 A; 9367 C; 9223 G; 13362 T; 0 U; 0 Other;

Query Match 6.7%; Score 35.6; DB 4; Length 48727;
Best Local Similarity 53.6%; Pred. No. 21;
Matches 74; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 246 CTAGAAACCCAAAGATTGTACCTTCGAATGTGATACACAGATTGCTTATTCAAATGTC 305
DB 3174 CTTACAGACCTAAGAGATTGCTTCTTGGAAATAAATACAGGTTTGTAGTATTTTAGATGAC 3115
QY 306 CTCGTTTCAGAGGAAAGATGATAATTAACCCCACTGCACCACTTGTTCACCGGCT 365
DB 3114 CTAATTAAAGTATACCACTTATGAGAGAGGCTTGTGAAGCGCTTGTGATGATGCTT 3055
QY 366 ATCAGGGTTGCTACTATT 383
DB 3054 TTAAGTTTACACCTTTT 3037

RESULT 11
ADA13316/C
ID ADA13316 standard; DNA; 108359 BP.
XX
AC ADA13316;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human fringe-like secreted protein gene, SEQ ID NO:3.
XX
KW Human; secreted protein; glycosyltransferase family; fringe protein;
KW testis; hepatocellular carcinoma; placenta; germinal centre B cell;
KW brain; chromosome 13; drug screening; cancer; cytostatic; gene therapy;
KW biochip; transgenic animal; gene; ds.
XX
OS Homo sapiens.
XX
XX
PH Key Location/Qualifiers
FT variation replace(2220,A)
FT /*tag= a
FT /note= "Single nucleotide polymorphism (SNP)"
FT CDS 2999..103512
FT /*tag= c
FT /partial
FT /product= "Human fringe-like secreted protein"
FT /note= "Encodes residues 53-499 of human fringe-like
FT secreted protein; no start codon given; contains 11
FT introns"
FT 2999..3112
FT /*tag= b
FT /transl_except= (pos:2999..3001;aa:Ile)
FT /transl_except= (pos:3101..3103;aa:His)
FT 3113..30842
FT /*tag= d
FT variation replace(3639,G)
FT /*tag= e
FT /note= "Single nucleotide polymorphism (SNP)"
FT variation replace(3722,A)
FT /*tag= f
FT /note= "Single nucleotide polymorphism (SNP)"
FT variation replace(4673,G)
FT /*tag= g
FT /note= "Single nucleotide polymorphism (SNP)"
FT variation replace(6235..6237,Gr)
FT /*tag= h
FT /note= "Single nucleotide polymorphism (SNP)"
FT variation replace(11368,T)
FT /*tag= i
FT /note= "Single nucleotide polymorphism (SNP)"
FT variation replace(12969,T)

FT	/*tag= j	replace(13998,C)	variation	FT	replace(31202,A)
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= k	FT	/*tag= aj	/*note= "Single nucleotide polymorphism (SNP) "
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= l	FT	replace(31739,G)	/*tag= ak
FT	replace(14144,A)	/*tag= m	FT	/note= "Single nucleotide polymorphism (SNP) "	replace(31878,G)
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= n	FT	/*tag= al	/*tag= al
FT	replace(14280,G)	/*tag= o	FT	/note= "Single nucleotide polymorphism (SNP) "	replace(32114,C)
FT	/*tag= p	/*tag= p	FT	/*tag= am	/*tag= am
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= q	FT	replace(33263,G)	/*tag= an
FT	replace(14294. .14296,TC)	/*tag= r	FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= ao
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= s	FT	replace(34267,G)	/*tag= ap
FT	replace(14309. .14311,TA)	/*tag= t	FT	/note= "Single nucleotide polymorphism (SNP) "	34768. .34904
FT	/*tag= u	/*tag= v	FT	/*tag= ar	/*tag= ap
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= w	FT	replace(34953. .34955,AT)	34905. .43035
FT	replace(16451,A)	/*tag= x	FT	/*tag= as	/*tag= aq
FT	/*tag= D	/*tag= y	FT	/note= "Single nucleotide polymorphism (SNP) "	replace(35080,A)
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= z	FT	replace(35112,G)	/*tag= at
FT	replace(18567,A)	/*tag= aa	FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= au
FT	/*tag= q	/*tag= ab	FT	replace(35987,G)	/*tag= av
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= ac	FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= ax
FT	replace(19008,T)	/*tag= ad	FT	replace(36052,T)	/*tag= ay
FT	/*tag= r	/*tag= ae	FT	replace(36466,T)	/*tag= az
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= af	FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= ba
FT	replace(19050,T)	/*tag= ag	FT	replace(37241,A)	/*tag= bb
FT	/*tag= s	/*tag= ah	FT	/*tag= bc	/*tag= bc
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= ai	FT	replace(37852,A)	/*tag= bd
FT	replace(19074,T)	/*tag= aj	FT	/*tag= be	/*tag= be
FT	/*tag= t	/*tag= ak	FT	43036. .43099	43100. .48332
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= al	FT	/*tag= bf	/*tag= bf
FT	replace(20487. .20489,CG)	/*tag= am	FT	replace(43425,A)	/*tag= bg
FT	/*tag= u	/*tag= an	FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= bh
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= ao	FT	replace(43806,T)	/*tag= bi
FT	replace(20776,C)	/*tag= ap	FT	/*tag= bh	/*tag= bi
FT	/*tag= v	/*tag= aq	FT	48333. .48452	/*tag= bi
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= ar	FT	/*tag= bi	
FT	20843. .20919	/*tag= as	FT		
FT	/*tag= w	/*tag= at	FT		
FT	20920. .21674	/*tag= au	FT		
FT	/*tag= x	/*tag= av	FT		
FT	replace(20923,T)	/*tag= ax	FT		
FT	/*tag= y	/*tag= ay	FT		
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= az	FT		
FT	replace(21188,A)	/*tag= ba	FT		
FT	/*tag= z	/*tag= bb	FT		
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= bc	FT		
FT	21675. .21786	/*tag= bd	FT		
FT	/*tag= aa	/*tag= be	FT		
FT	21787. .34767	/*tag= bf	FT		
FT	/*tag= ab	/*tag= bg	FT		
FT	replace(22922,A)	/*tag= bh	FT		
FT	/*tag= ac	/*tag= bi	FT		
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= bj	FT		
FT	replace(23240. .23242,TC)	/*tag= bk	FT		
FT	/*tag= ad	/*tag= bl	FT		
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= bm	FT		
FT	replace(23272,G)	/*tag= bn	FT		
FT	/*tag= ae	/*tag= bo	FT		
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= bp	FT		
FT	replace(23753,A)	/*tag= bq	FT		
FT	/*tag= af	/*tag= br	FT		
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= bs	FT		
FT	replace(24309,G)	/*tag= bt	FT		
FT	/*tag= ag	/*tag= bu	FT		
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= bv	FT		
FT	replace(24333,T)	/*tag= bw	FT		
FT	/*tag= ah	/*tag= bx	FT		
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= by	FT		
FT	replace(24505,A)	/*tag= bz	FT		
FT	/*tag= ai	/*tag= ca	FT		
FT	/note= "Single nucleotide polymorphism (SNP) "	/*tag= cb	FT		

```
FT intron 48453. .50526
FT /*tag= bj
FT variation replace(49219,C)
FT /*tag= bk
FT /note= "Single nucleotide polymorphism (SNP)"
FT replace(49259,A)

Query Match 6.7%; Score 35.6; DB 9; Length 108359;
Best Local Similarity 45.9%; Pred. No. 30;
Matches 122; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 235 AGAATCGACCCCTAGAAACCCAAAGATTGTGACCTTCGAATGTGATACACAGATTGCTTA 294
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
850 ACAATATAAACATATAATATAATATAATATAATATAATATAATATAATATAATATAATATA 791
QY 295 TTCAAAATGTCCTCGTTCAGAGGAAGATGATAATTAACCCACTGGATGCCACCTTG 354
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
790 TGAAGGTATAAAGATTACGTGAGTATGATAAATACCAATCAGGATAATGGTTAAAGG 731
QY 355 TTGCACGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGATTTTCTGCTGAAGG 414
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
730 TTTCAAGTGTGTTAATGGTATTAAATTTCTTAATCTAGGTGATATTTTCTTCTTAATCT 671
QY 415 AGAGAGTCTGGAACCCAGACCACTGCTTATTTCTTAATCAATCATATGTTGTTATCTATC 474
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
670 AATCAACATATAATACATGAATGTTCAATATATATTCTCTTATAACTTCTTATATGTT 611
QY 475 AAAAAAATATGTATGCATGATATA 500
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
610 CAAATATTTCCTAACTACTTTTTTTA 585

RESULT 12
AAS46563/C
ID AAS46563 standard; DNA; 6134 BP.
XX AC AAS46563;
XX DT 18-DEC-2001 (first entry)
XX DE Tumour suppressor gene derived chemically modified sequence #285.
XX KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;
XX KW tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
XX KW cytosine methylation; ds.
XX OS Homo sapiens.
XX PN WO200168912-A2.
XX PD 20-SEP-2001.
XX PF 15-MAR-2001; 2001WO-EP002955.
XX PR 15-MAR-2000; 2000DE-01013847.
XX PR 06-APR-2000; 2000DE-01019058.
XX PR 07-APR-2000; 2000DE-01019173.
XX PR 30-JUN-2000; 2000DE-01032529.
XX PR 01-SEP-2000; 2000DE-01043826.
XX PA (BPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX XX WPI; 2001-602752/68.
XX XX Fragments of chemically modified genes associated with tumor suppressor
XX PT genes and oncogenes, useful in designing primers and probes for analyzing
XX PT diseases associated with cytosine methylation state e.g. cancer.
XX PS Claim 1; SEQ ID NO 285; 27pp; English.
XX XX The invention relates to a nucleic acid comprising a sequence of 18
```

```
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and oncogenes
CC having a sequence taken from 536 (actually 533 since numbers 408, 458 and
CC 500 are missing from the sequence listing) sequences (Ss) and sequences
CC complementary to (Ss). The nucleic acid may be a peptide nucleic acid-
CC oligomer (PNA) of at least 9 nucleotides and may form part of a set of
CC probes for detecting the cytosine methylation state and/or single
CC nucleotide polymorphisms and also to be used in an array for analysing
CC diseases associated with CpG dinucleotides e.g. cancers and tumours. The
CC probes can also be used in a method for ascertaining genetic and/or
CC epigenetic parameters for the diagnosis and/or therapy of existing
CC diseases or the predisposition to specific diseases, by analysing
CC cytosine methylations. The parameters may be compared to another set of
CC genetic and/or epigenetic parameters, the differences serving as basis
CC for diagnosis and/or prognosis events which are disadvantageous to
CC patients. The present sequence is one of the 533 genomic sequences
CC derived from tumour suppressor genes and oncogenes. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 6134 BP; 1576 A; 76 C; 1372 G; 3110 T; 0 U; 0 Other;

Query Match 6.7%; Score 35.2; DB 4; Length 6134;
Best Local Similarity 44.9%; Pred. No. 12;
Matches 133; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 220 TTTGATTTGTGAGGAGAAATCTGACCCCTAGAACCCCAAGATTGTACCTTCGAATGTGA 279
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4083 TTCGAAATTTTACAAAAATAAATAAACATTAACATAAATAATAATAATAATAATAATAA 4024
QY 280 TACACAGATTGCTTATTCAAATGTCCTCGTTTCAGAGGAAGATGATAATTTAAACCCAC 339
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4023 ATATAAAATTACATCTACAAATATAACCTCAATACAAAAAATAAATAAATAAATAAAT 3964
QY 340 TGGATGCACCACTTGTTCGACGGGCTATCAGGGTTCCTACTATTTTCGATCAAGATGTGA 399
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3963 AAAAAATATATTAATAAAAAACAAAAATAATACTATTATCTTTAATAATATATCAAAATTTTC 3904
QY 400 TTTTGTCTGTGAGGAGAGAGTCTCGAACCCCAAGACCACTGCTTATTCTTAATCAATCAT 459
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3903 CTTTCTCTATTTTTCATTAACCTCTTAACACAAAAACCAATATATTTATATAAAAAAA 3844
QY 460 ATGTTGTTATCTATCAAAAAAATAATATGATGATGATATATGCTGTTACTGTAA 515
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3843 AATAATTTATTTTAACTAAACACATAACTCATCTATATAATCCCAAAACTTTTA 3788

RESULT 13
ABL32638
ID ABL32638 standard; DNA; 6650 BP.
XX AC ABL32638;
XX XX 26-MAR-2002 (first entry)
XX DT Human immune system associated gene SEQ ID NO: 611.
XX DE Human; immune system disease; cytosine methylation; antiasthmatic;
XX KW antiarteriosclerotic; anti-naemic; cytosine; cytosine; cytosine;
XX KW neuroprotective; anti-HIV; anti-naemic; cytosine; cytosine;
XX KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
XX KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
XX KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
XX KW ds.
XX OS Homo sapiens.
XX XX WO200200928-A2.
XX PN 03-JAN-2002.
XX PD
```


XX Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
XX
PS Claim 1; SEQ ID NO 71; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SQ Sequence 9964 BP; 2847 A; 42 C; 1995 G; 5080 T; 0 U; 0 Other;

Query Match 6.6%; Score 35; DB 6; Length 9964;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 50; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 455 ATCATATGTTGTTATCTATCAAAAAAATATGATGCATGATATATGCTGTTACTGTA 514
Db 3825 ATTTAGGATGTTATTTTAAAAAAATGTTTTATAAATGTTAAAAAGATGTTTATTGTA 3884

Qy 515 ATGTGGACTTTATTG 529
Db 3885 ATATTAAATTATATG 3899

Search completed: January 22, 2005, 05:47:32
Job time : 414 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 01:01:52 ; Search time 2760 Seconds

(without alignments)
9063.859 Million cell updates/sec

Title: US-10-725-829-1

Perfect score: 529

Sequence: 1 cataatggctgtcacaaag.....ctgtaatgtgactttatttg 529

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_cm.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	529	100.0	529	8	AF174381 Solanum a
2	288	54.4	687	8	AY422686 Solanum n
3	288	54.4	766	8	LECSVI157G
4	279.6	52.9	838	8	X94946 L.esculentu
5	265.4	50.2	660	8	STPIN2W
6	265.4	50.2	1776	8	K03291 Tomato leaf
7	263.6	49.8	1695	8	BT013250 Lycopersi
8	260.6	49.3	2068	8	STPIN2G
9	259.2	49.0	559	8	STU45450
10	259.2	49.0	684	8	AB110700 Lycopersi
11	259.2	49.0	1670	8	AY007240 Lycopersi
12	258	48.8	1241	8	AY129402 Lycopersi
13	257.6	48.7	482	8	M5186 S.tuberosum
14	256.8	48.5	512	8	AY247794 Solanum p
15	256.8	48.5	666	8	X03778 Potato (Sol
16	256.4	48.5	1274	8	L37519 Solanum tub
17	255.8	48.4	2330	8	Z29537 N.tabacum (
18	254.2	48.1	584	8	Z13992 S.tuberosum
19	251.8	47.6	580	8	X78275 S.tuberosum
					AY517498 Solanum p

ORIGIN

20	250.2	47.3	1914	8	STPI2G
21	249	47.1	554	8	STPIIR2
22	227.4	43.0	532	8	AF209709
23	219.4	41.5	1573	8	STPRINPSG
24	178.8	33.8	1342	8	NAU08219
25	178.8	33.8	1360	6	E54395
26	178.8	33.8	1360	6	AR224422
27	178.8	33.8	1360	6	AR230263
28	175.6	33.2	946	8	AF105340
29	164.4	31.1	1546	8	AF542547
30	158	29.9	1795	8	AF205851
31	149.6	28.3	764	8	AF221097
32	149	28.2	1104	6	E54394
33	149	28.2	1104	6	AR224421
34	149	28.2	1104	6	AR230262
35	148.8	28.1	1329	8	AF205852
36	143.8	27.2	836	8	AF039398
37	132.8	25.1	1414	8	AY426751
38	114.2	21.6	573	8	AF242734
39	105.8	20.0	562	8	BT013127
40	93	17.6	841	8	TOMARPIX
41	87.8	16.6	4272	8	TOMARPI
42	66.2	12.5	175	8	AY184823
43	42.8	8.1	1379	5	CR733321
44	42.2	8.0	536	11	G58878
45	42	7.9	235502	2	AC103222

ALIGNMENTS

RESULT 1	AF174381	529 bp	mRNA	linear	PLN 15-JAN-2002
LOCUS	Solanum americanum	proteinase inhibitor IIA (PIN2a)	mrna	complete	
DEFINITION	Solanum americanum	proteinase inhibitor IIA (PIN2a)	mrna	complete	
ACCESSION	AF174381				
VERSION	AF174381.1	GI:17221674			
KEYWORDS	Solanum americanum				
SOURCE	Solanum americanum				
ORGANISM	Solanum americanum				
REFERENCE	Xu, Z.F., Qi, W.Q., Ouyang, X.Z., Yeung, B. and Chye, M.L.				
AUTHORS	Xu, Z.F., Qi, W.Q., Ouyang, X.Z., Yeung, B. and Chye, M.L.				
TITLE	A proteinase inhibitor II of Solanum americanum is expressed in				
JOURNAL	Phloem				
MEDLINE	Plant Mol. Biol. 47 (6), 727-738 (2001)				
PUBMED	21643209				
REFERENCE	2 (bases 1 to 529)				
AUTHORS	Chye M.-L. and Xu Z.-F.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-AUG-1999) Department of Botany, The University of				
FEATURES	Hong Kong, Pokfulam Road, Hong Kong, China				
source	Location/Qualifiers				
	1..529				
	/organism="Solanum americanum"				
	/mol_type="mRNA"				
	/db_xref="taxon:109975"				
gene	1..529				
	/gene="PIN2a"				
CDS	5..451				
	/gene="PIN2a"				
	/codon_start=1				
	/product="proteinase inhibitor IIA"				
	/protein_id="AAL36458.1"				
	/db_xref="GI:17221675"				
	/translation="MAVHKVSFLACLLVLGMFLAKHVDKACTRECGRHSGYICPR				
	SGSGQKPICTNCCSGYKCNYSAGDLICEGSDPRNPDKDCTFECDTQIAYSKCPR				
	SEGRMKIRPTGCTTCTCTGCGCYFYDDQDGFVCEGSESPFKTAYF"				

JOURNAL Submitted (10-JAN-1996) P. Vera, Universidad Politecnica de Valencia, Inst. de Biologia Molecular y Celular de Plantas, Camino de Vera 14, E- 46022 Valencia, SPAIN

FEATURES
source
1. .766
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="Rutgers"
/db_xref="taxon:4081"
/tissue_type="leaf"
1. .766
/genes="cevi57"
43. .548
/genes="cevi57"
/codon_start=1
/product="proteinase inhibitor II"
/protein_id="CAA64416.1"
/db_xref="GI:1161572"
/db_xref="GOA:Q43502"
/db_xref="Swiss-Prot:Q43502"
/translation="MAVYKVSFLAHLLVLMYLLSTVEHANACTKECNLGVGICPG SEGSNPICITNCSSGKGCNYYANGTFCIGTSDPKNPICPSYQDIAYSKCPR SEGKTIYPTGCTCTCYGKCYFQDGEFVCEGSIIEPKGCTKCDPRVMTCP SGLAKLNOVCNCCSAGEGCKLYDNDGSLLTGEPQSI STA"
43. .111
/genes="cevi57"

sig_peptide
43. .111
/genes="cevi57"

ORIGIN
Query Match 54.4%; Score 288; DB 8; Length 766;
Best Local Similarity 79.2%; Pred. No. 3.6e-66;
Matches 342; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
QY 2 ATAATGGCTGTTCCAAAAGTTAGCTTCCTTGGCTTGCTACTTCTTGGATGGATGTTT 61
DB 40 ACAATGGCTGTTTACAAAAGTTAGTTCTTGGCTCACCTACTTCTTGGAAATGATCTA 99
QY 62 CTACTTGCAGAACATGTTGATGCCAAGCTTGTACTAGAGAAATGTGGTCAATTTAGCTAT 121
DB 100 CTAGTAAGCACGGTGGAAACACGCTAATGCTTGTACTAAAGAAATGTGGTAACTTGGCTAT 159
QY 122 GGCATATGCCACGTTTCAGAAAGGAAGTCCCAAAAACCTATATGCACCAATGTGTGCTCA 181
DB 160 GGGATATGCCACGTTTCAGAAAGGAAGTCCAGAAATCCATATGTACCAATTTGTGCTCT 219
QY 182 GGCATATAAGGGTTGCAACTATTACAGTGTCTAAAGGAGATTGATTTGTGAAGGAGATCT 241
DB 220 GGCTATAAGGGTTGCAACTATTATACGCTAATGGAACCTTTTATTTGTGAAGGAACGTCT 279
QY 242 GACCCTAGAAACCCAAAGATGTGACCTCGAATGTGATACACAGATGCTTATTCAAA 301
DB 280 GATCCAAAATCTTAACATTTGCCCTCATATTGTGATCCAAAATTTGCTATTCAAAG 339
QY 302 TGTCCCTGTTTCAGAAAGAAAGATGATAATTAACCCACCTGGATGCACCACTGTGTGCAG 361
DB 340 TGTCCAGTTTCAGAAAGAAAGAGATATCTATCCACAGATGTGACGCTGTGCAC 399
QY 362 GGCTATCAGGGTTGCTACTATTTTCGATCAAGATGGTGAATTTGTCTGTGAAGGAGAGAT 421
DB 400 GGTTCAGAAAGGGTTGCTACTATTTTGGTCAAGATGGAGAGTTTGTGTGAAGGAGAGAT 459
QY 422 CCTGAACCCCAAG 433
DB 460 ATTGAACCTAAG 471

STPIN2W 838 bp mRNA linear PLN 05-AUG-1997
LOCUS S.tuberosum mRNA for pin2 gene, wound induced.
DEFINITION X99095
ACCESSION X99095.1 GI:1431616
VERSION pin2 gene.
KEYWORDS Solanum tuberosum (potato)

ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE
AUTHORS Damann,C., Rojo,E. and Sanchez-Serrano,J.J.
TITLE Abscissic acid and jasmonic acid activate wound-inducible genes in potato through separate organ-specific signal transduction pathways
JOURNAL Plant J. 11, 101-110 (1997)
REFERENCE 2 (bases 1 to 838)
AUTHORS Damann,C.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-1996) C. Damann, Centro Nacional de Biotecnologia, Plant Genetics, Campus de Cantoblanco, 28049 Madrid, Spain

FEATURES
source
1. .838
/organism="Solanum tuberosum"
/macronuclear
/mol_type="mRNA"
/sub_species="Desiree"
/db_xref="taxon:4113"
/haplotype="tetraploid"
/cell_line="Desiree"
/tissue_type="root"
1. .838
/gene="pin2"
19. .111
/gene="pin2"
/evidence=not_experimental

sig_peptide
19. .111
/gene="pin2"

ORIGIN
Query Match 52.9%; Score 279.6; DB 8; Length 838;
Best Local Similarity 80.2%; Pred. No. 6.4e-64;
Matches 356; Conservative 0; Mismatches 79; Indels 9; Gaps 2;
QY 1 CATATGGCTGTTCACA---AAGTTAGCTTCCTTGGCTTGCCTACTTGTCTTGGATGGAT 57
DB 15 CATATGGCTGTTCACAAAAGAAAGTTAGTTCCCTTGTCTTACCTACTTGTCTTGGATTA 74
QY 58 GTTTCCTACTTGC-----GAAACATGTTGATGCAAGGCTTGTACTAGAGAATGTGTCTCA 111
DB 75 GGTACTTGTAAAGCGCGATGGACATGTTGATGCGAAGGCTTGTACTTTAGAAATGGTAA 134
QY 112 TTTTACTATGGCATATGCCACGTTTCAGAAAGAACTCCCAAAAACCTATATATGCACAA 171
DB 135 TCTTGGCTATGGATATGCCACGTTTCAGAAAGAACTCCCAAAAATCCCATATATGCACAA 194
QY 172 TTGTTGCTCAGGCTATAAGGTTGCAACTATTACAGTGTCTTAAGGAGATTTGATTTGTGA 231
DB 195 CTGTTGTGAGGTTATAAAGGTTGCAATTTATATAGTGTCTTAATGGGACTTTTCATTTGCGA 254
QY 232 AGGAGAACTCTGACCTCTAGAAACCCAAAAGATTTGACTCTCGAATGTGATACACAGATTGC 291
DB 255 AGGACATCTGACCCAAAACCCAAAAGCTTGGCCCGGAAATTTGATCCACATATTGC 314
QY 292 TTATTCAAATGTCTCTGTTTCAGAAAGAAAGATGATAATTAACCCCACTGGATGCACAC 351
DB 315 CTATTCAAATGTCTCCGTTTCAAGAGGAAAGACACTAAATTTATCCCAACAGATGTACCA 374
QY 352 TTGTTGCAGGGCTATCAGGTTGCTACTATTTCATTCAGATGGTGTGATTTGCTCTGTGA 411
DB 375 ATGTTGCAGGGGTACAAAGGTTGCTACTATTTTGTGTAAGACGGCAAGTTTGTGTTGTGA 434
QY 412 AGGAGAGAGTCTGAAACCCCAAGAC 435
DB 435 AGGAGAGATTTTGAACCCCAAGGC 458

RESULT 5
LOCUS TOMWIP11
DEFINITION Tomato leaf wound-induced proteinase inhibitor II mRNA, complete

```

cds.
K03291
K03291.1 GI:170521
protease inhibitor; wound-induced protease inhibitor.
Lycopersicon esculentum (tomato)
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 660)
Graham, J.S., Pearce, G., Merryweather, J., Titani, K., Ericsson, L.H.
and Ryan, C.A.
Wound-induced proteinase inhibitors from tomato leaves. II. The
cDNA-deduced primary structure of pre-inhibitor II
J. Biol. Chem. 260 (11), 6561-6564 (1985)
85207658
3838986
Original source text: Tomato (L. esculentum L. var. Bonnie Best)
wounded leaf, cDNA to mRNA, clone pTI-24.
A 10 bp palindrome identical to the one following the TOMWPII cds
(see separate entry) begins at position 586 [1]. A polyadenylation
signal is present at position 638-643 [1].
A printed copy of this sequence was kindly provided by C.A. Ryan
(18-SEP-1985).
FEATURES
source
Location/Qualifiers
1..660
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
<1..660
/product="WPII mRNA"
49..495
/notes="wound-induced proteinase inhibitor II prepeptide"
/codon_start=1
/protein_id="AAA34201.1"
/db_xref="GI:170522"
/translation="MAVHKEVNFVAYLLVLGMLFYDADACTRECNGLFGICPRSE
GSLPINCISGSGYKGCNYNSFGKFIGESDPRKPNACTFNCDPNIAYSRCPRSQ
GKSLIYPTGCTCTCTGKGYGKGYFGKDGKFCGESDEPRKNMYPVM"
49..123
sig_peptide
/notes="wound-induced proteinase inhibitor II signal
peptide"
124..492
mat_peptide
/product="wound-induced proteinase inhibitor II mature
peptide"
279 bp upstream of Sau3A site.
ORIGIN
1 CATAATGGCTGTTCACA---AAGTTAGCTTCCTTGGCTTACTTGTCTTGGATGGAT 57
Best Local Similarity 50.2%; Score 265.4; DB 8; Length 660;
Matches 349; Conservative 0; Mismatches 86; Indels 12; Gaps 2;
QY 1 CATAATGGCTGTTCACA---AAGTTAGCTTCCTTGGCTTACTTGTCTTGGATGGAT 57
DB 45 CATCATGGCTGTTCACAAGGAAGTTAATTTGTGCGCTTACCTACTAATTTCTTGGAT 104
QY 58 GTTCTTACTTGGCAACATGTTGATGCGCAAGGCTTGCTAGAGAAATGGTCAATTTAG 117
DB 105 GTTTCT-----ATATGTTGATGCCAAGGCTTGCTAGAGAAATGGTCAATTTAG 155
QY 118 CTATGGCATATGCCACAGTTTCAAGGAAGTCCCAAAAAACCTATATGCAACCAATTTGG 177
DB 156 GTTCGGGATATGCCACAGTTTCAAGGAAGTCCGCTAAATCCCATATGCAATCAATTTGG 215
QY 178 CTCAGCTATAAGGGTTGCAACTTATACAGTGTCTAAAGGAGATTGATTTGTGAAGAGA 237
DB 216 CTCAGGCTATAAGGGTTGTAATTTATATAATTTCTTCGGAATAATTTATTTGTGAAGAGA 275
QY 238 ATCTGACCTTAGAAACCCAAAGATTGTACCTTCGAATGTATGATACACAGATTCCTTATTC 297
DB 276 ATCTGATCCAAAAGGCCAAATGCTTGCACCTTTAATTTGTGATCCAAATATTCCTTATTC 335
QY 298 AAAATGCTCTGTTTCAGAAGGAAGATGATAATTAACCCACCTGGATGCAACCAATTTGG 357
QY 358 CACGGCTATACAGGGTTGCAACTTATACAGTGTCTAAAGGAGATTGATTTGTGAAGAGA 417
DB 451 GAGTGTGATGAACCCCAAGGCAAAATATGTA 477

```

```

Db 336 AAGATGTCCCGTTTCAAGGAAAATCGTTAATTTATCCACCGGATGACACGTTGTTG 395
QY 358 CACGGCTATACAGGGTTGCTACTATTTCGATCAAGATGGTGAATTTGTCTGTGAAGAGA 417
Db 396 CACAGGGTACAAGGGTTGCTATTATTATTTGGTAAAGATGGAAGCTTTGTATGTGAAGAGA 455
QY 418 GAGTCTCTGAACCCCAAGACCACTGCTTA 444
Db 456 GAGTGTGATGAACCCCAAGGCAAAATATGTA 482

RESULT 6
BT013250
LOCUS
DEFINITION
Lycopersicon esculentum clone 134756F, mRNA sequence.
BT013250
ACCESSION
BT013250.1 GI:47104665
FLI_CDNA.
Lycopersicon esculentum (tomato)
ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1776)
Kirkness, E.F., Wang, W. and Vazeille, A.
Direct Submission
Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
Medical Center Drive, Rockville, MD 20850, USA
FEATURES
source
Location/Qualifiers
1..1776
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="134756F"
/tissue_type="mixed elicitor"
/note="TMECD39"

ORIGIN
Query Match 50.2%; Score 265.4; DB 8; Length 1776;
Best Local Similarity 78.1%; Pred. No. 3.9e-60;
Matches 349; Conservative 0; Mismatches 86; Indels 12; Gaps 2;
QY 1 CATAATGGCTGTTCACA---AAGTTAGCTTCCTTGGCTTACTTGTCTTGGATGGAT 57
DB 40 CATCATGGCTGTTCACAAGGAAGTTAATTTGTGCGCTTACCTACTAATTTCTTGGAT 99
QY 58 GTTCTTACTTGGCAACATGTTGATGCGCAAGGCTTGCTAGAGAAATGGTCAATTTAG 117
DB 100 GTTTCT-----ATATGTTGATGCCAAGGCTTGCTAGAGAAATGGTCAATTTAG 150
QY 118 CTATGGCATATGCCACAGTTTCAAGGAAGTCCCAAAAAACCTATATGCAACCAATTTGG 177
DB 151 GTTCGGGATATGCCACAGTTTCAAGGAAGTCCGCTAAATCCCATATGCAATCAATTTGG 210
QY 178 CTCAGGCTATAAGGGTTGCAACTTATACAGTGTCTAAAGGAGATTGATTTGTGAAGAGA 237
DB 211 CTCAGGCTATAAGGGTTGTAATTTATATAATTTCTTCGGAATAATTTATTTGTGAAGAGA 270
QY 238 ATCTGACCTTAGAAACCCAAAGATTGTACCTTCGAATGTATGATACACAGATTCCTTATTC 297
DB 271 ATCTGATCCAAAAGGCCAAATGCTTGCACCTTTAATTTGTGATCCAAATATTCCTTATTC 330
QY 298 AAAATGCTCTGTTTCAGAAGGAAGATGATAATTAACCCACCTGGATGCAACCAATTTGG 357
DB 331 AAGATGTCCTGTTTCAAGGAAAATCGTTAATTTATCCACCGGATGTACCAATTTGG 390
QY 358 CACGGCTATACAGGGTTGCTACTATTTCGATCAAGATGGTGAATTTGTCTGTGAAGAGA 417
DB 391 CACAGGGTACAAGGGTTGCTATTATTATTTGGTAAAGATGGAAGCTTTGTATGTGAAGAGA 450
QY 418 GAGTCTCTGAACCCCAAGACCACTGCTTA 444
DB 451 GAGTGTGATGAACCCCAAGGCAAAATATGTA 477

```



```

RESULT 7
STRIN2G
LOCUS STPRIN2G 1695 bp DNA linear PLN 11-MAY-1995
DEFINITION S. tuberosum gene for proteinase inhibitor II.
ACCESSION 212753
VERSION 212753.1 GI:21553
KEYWORDS proteinase inhibitor; proteinase inhibitor II.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
AUTHORS Choi,Y., Moon,Y. and Lee,J.S.
TITLE Primary structure of two proteinase inhibitor II genes closely
linked in the Potato genome
JOURNAL Hanguk Saenghwahakhoe Chi 23, 214-220 (1990)
REFERENCE
AUTHORS Lee,J.S.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-1992) Jong S. Lee, Molecular Biology, Seoul
National University, ShinRim Dong, Kwanak-Gu, Seoul, 151-742,
Republic of Korea
FEATURES
source
1. .1695
/organism="Solanum tuberosum"
/mol_type="genomic DNA"
/cultivar="Russel Burbank"
/db_xref="taxon:4113"
repeat_region
40..148
/notes="partial long direct repeat unit 1"
repeat_region
262..378
/notes="partial long direct repeat unit 2"
CAAT_signal
438..441
TATA_signal
545..550
CDS
join(629..677,779..1206)
/codon_start=1
/product="proteinase inhibitor II"
/protein_id="CAA78277.1"
/db_xref="GI:21554"
/db_xref="GOA:Q00782"
/db_xref="Swiss-Prot:Q00782"
/translation="MAHKVSVFLAYLLVGLMLFVSMHVDKACTLECNLGYGI
CPRSEGPENICNCAGYKGNYSANGTFICEGQSHPKPKACPRNCDPHIAYSK
CPRSGGKTLIYPTGCTTCCTGYTDCYVFGKDGKFCVCEGSEIEPKACTLECDRLHT"
629..677
/number=1
exon
678..778
intron
779..1206
/number=2
polyA_signal
1476..1481
misc_structure
1562..1568
/notes="mRNA 3' maturation signal"
ORIGIN
Query Match 49.8%; Score 263.6; DB 8; Length 1695;
Best Local Similarity 82.5%; Pred. No. 1.2e-59;
Matches 302; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 70 GAAACATGTTGATGCCAAGCTTGTTACTAGAGAATGGTCATTTTAGCTATGGCATATG 129
Db 804 GGAACATGTTGATGCCAAGCTTGTTACTTTAGAATGGTAACTTTGGCTATGGCATATG 863
QY 130 CCACGTTACAGAGGAGTCCCAAAACCTATATGCACCAATGTTGCTCAGGCTATAA 189
Db 864 CCCACGTTACAGAGGAGTCCGAAAATCCCATATGCACCACTGTTGTGCGGTTATAA 923
QY 190 GGGTTGCAACTATTACAGTCTCTAAAGGAGATTTGATTGTGAAGGAGAACTCTGACCTAG 249
Db 924 AGTTGCAACTATTATAGTCTCTAATGGGACCTTTTATTTTGAAGGACAATCTCACCACA 983

```

```

QY 250 AAACCCAAAAGATTGTACCTTCGAATGTGATACACAGATTGCTTATTCAAATGTCTCTCG 309
Db 984 AAACCCAAAAGCTTGCCTCCCGAAATTTGTATCCACATATTGCTTATTCAAATGTCCACG 1043
QY 310 TTCAAGAGAAAGATGATAATTAACCCACTGTGATGCACCACTTGTTCACGGGCTATCA 369
Db 1044 TTCAGGAGAAAGACCTTAATTTATCCACAGGATGTACCACATGTTGCACGGGATACAC 1103
QY 370 GGGTTCTCTACTATTTCGATCAAGATGGTGAATTTTGTCTGTGAAGGAGAGAGTCTGAACC 429
Db 1104 GGATTCTCTATTATTTTGGTAAAGACGGCAAGTTTGTGTGTGAAGGAGAGAGTATTGAACC 1163
QY 430 CAAGAC 435
Db 1164 CAAGGC 1169

STU45450 2068 bp DNA linear PLN 27-JAN-1999
LOCUS Solanum tuberosum proteinase inhibitor II (pin2r) gene, complete cds.
DEFINITION STU45450
ACCESSION U45450
VERSION U45450.1 GI:1206014
KEYWORDS Solanum tuberosum (potato)
SOURCE Solanum tuberosum
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE
AUTHORS Park,S. and Thornburg,R.
TITLE Isolation and characterization of a proteinase inhibitor II gene
that is not wound-inducible (Accession No. U45450) (PGR96-007)
JOURNAL Plant Physiol. 110, 1048 (1996)
REFERENCE
AUTHORS Thornburg,R.W. and Park,S.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-1996) Robert W. Thornburg, Biochemistry and
Biophysics, Iowa State University, Ames, IA 50011, USA
FEATURES
source
1. .2068
/organism="Solanum tuberosum"
/mol_type="genomic DNA"
/strain="Russel Burbank"
/db_xref="taxon:4113"
1014..1917
/gene="pin2"
join(1014..1109,1223..1917)
/gene="pin2"
join(1058..1109,1223..1614)
/gene="pin2"
/function="inhibitor of trypsin and chymotrypsin"
/codon_start=1
/product="proteinase inhibitor II"
/protein_id="AAD09849.1"
/db_xref="GI:1206015"
/translation="MAVHKVSVFVAYLLVGLMFLYVDALGCTKECNLGFICPRSE
GSPTNPICNCCSGYKGNYSAGFLICQESDPKNPKACPLNCDTATNAYSRCPRSE
GKSLIYPTGCTTCCTGCTGKCYGTFGTNGKFCVCEGSEDEPKYMSA"
ORIGIN
Query Match 49.3%; Score 260.6; DB 8; Length 2068;
Best Local Similarity 82.4%; Pred. No. 7.4e-59;
Matches 299; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 73 ACATGTTGATGCCAAGCTTGTTACTAGAGAATGGTCAATTTTAGCTATGGCATATGCC 132
Db 1233 ATATGTTGATGCTTTGGTGTGTTACTAAAGAATGTGTTAACTTTGGCTTTGGGATATGCC 1292
QY 133 AGTTTCAGAGGAGTCCCAAAACCTATATGCACCAATTTGCTCAGGCTATAGGG 192
Db 1293 AGTTTCAGAGGAGTCCCAAAACCTATATGCACCAATTTGCTCAGGCTATAGGG 1352

```

```
QY 193 TTGCAACTATTACAGTCTAAAGGAGATTGATTTGTGAAGAGAATCTGACCCCTAGAAA 252
|||
Db 1353 TTGTAATTATTATAGTCTTCGGAGATTTAAATTTGCCAAGGAGAATCTGACCCAAAAA 1412
|||
QY 253 CCACAAAGATTGACCTTCCGAATGTGTATACACAGATGCTTATCAAAATGCTCGTTC 312
|||
Db 1413 CCCAAAGCTTGCCCCCTAAATTTGTGATACAAAATTTGCCCTATTCAAGATGCTCGTTC 1472
|||
QY 313 AGAAGGAAGATGATTAATTAACCCCACTGATGCACACATTTGTGCACGGGCTATCAGGG 372
|||
Db 1473 AGAAGGAATAATCGCTAATTTATCCCACTGGATGTACACATGCTGCACAGATACAGGG 1532
|||
QY 373 TTGCTACTATTTCGATCAAGATGGTGATTTTGTGTGAAGGAGAGAGTCTTGAACCCAA 432
|||
Db 1533 TTGCTACTATTTCGGTACAAATGCGAAGTTTGTGTATGTGAAGGAGAAAGTGATGAACCCAA 1592
|||
QY 433 GAC 435
|||
Db 1593 GCC 1595

RESULT 9
AB110700 559 bp DNA linear PLN 28-MAY-2003
LOCUS Lycopersicon esculentum gene for protease inhibitor II, complete cds.
DEFINITION
ACCESSION AB110700 GI:31088241
VERSION
KEYWORDS
SOURCE Lycopersicon esculentum (tomato)
ORGANISM
REFERENCE
1 Wada,M., Matsuda,Y., Fujita,K., Nanjo,A., Kakutani,K., Nonomura,T., Matsuda,Y. and Toyoda,H.
Mature mRNAs in targeted single cells can be amplified by RT-PCR of nucleus-free cytosolic contents sucked up with micropipette
Unpublished
2 (bases 1 to 559)
Nanjo,A., Wada,M., Fujita,K., Sameshima,T., Kakutani,K., Nonomura,T., Matsuda,Y. and Toyoda,H.
Direct Submission
Submitted (24-MAY-2003) Yoshinori Matsuda, Kinki University, Laboratory of Plant Pathology and Biotechnology Faculty of Agriculture; 3327-204 Nakamachi, Nara, Nara 631-8505, Japan (E-mail:ymatsuda@nara.kindai.ac.jp, Tel:81-742-43-1511(ex.3111), Fax:81-742-43-1155)
FEATURES
source Location/Qualifiers
1..559
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/cultivar="Money-maker"
/db_xref="taxon:4081"
join(2..53,165..559)
/codon_start=1
/product="protease inhibitor II"
/protein_id="BAC76901.1"
/db_xref="GI:31088242"
CDS
/translations="MAVHKVNFVAYLLIVLGMFLYVDKAKTREGCNLGFICPRSE
GSPLPNICNCSGYKCNYNYSFGKFIGESDPRKPNACTFNCDPNIAYSRCPRSQ
GKSLIYPTGCTTCTCTGYKGYFGDKGKFCVCESEDEPKANMYPVM"
ORIGIN
Query Match 49.0%; Score 259.2; DB 8; Length 559;
Best Local Similarity 74.3%; Pred. No. 1.9e-58;
Matches 327; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 5 ATGGCTGTTACAAAGTAGTCTTCCTGCTGCTACTGTTCTTGTGATGATGTTCTTA 64
|||
Db 107 ATTGGTTTATACATACACAGTAGTTTATATTATTTTCTTATATTATATTGTTGAGGA 166
|||
```

```
QY 65 CTTGCGAAACATGTTTGATGCCAAGGCTTGCTACTAGAGAATGTGGTCATTTTAGCTATGGC 124
|||
Db 167 ATGTTTCTATATGTTGATGCCAAGGCTTGCTACTAGAGAATGTGGTAATCTTGGGTTCCGG 226
|||
QY 125 ATATGCCACGTTTCAGAAGGAGTCCCAAAAACCTATATGACCACTATGTTGCTCAGGC 184
|||
Db 227 ATATGCCACGTTTCAGAAGGAGTCCCGCTAAATCCCATATGTCATCAATTTGTTCTCAGGC 286
|||
QY 185 TATAAGGGTTGCAACTATTACAGTGTCTAAGAGAGATTGATTTGTGAAGGAGAATCTGAC 244
|||
Db 287 TATAAGGGTTGTAATTTATTAATTTCTTCGGGAAAATTTATTGTAAGGAGAATCTGAT 346
|||
QY 245 CCTAGAAAACCCAAAGATTGTACCTTTCGAATGTGTATACACAGATTGCTTATTCAAAATGT 304
|||
Db 347 CCAAAAAGGCCAAATCGTTGCACCTTTAAATTTGTGATCCCAAATATTGCTTATTCAAGATGT 406
|||
QY 305 CCTCGTTTCAGAAGGAAAAGATGATAATAAACCCACGTGATGACCACTGTTGTCACGGGC 364
|||
Db 407 CCCCCTTTCACAAAGGAAAATCGTTAAATTTATCCACCGGATGTACCAATGTTGCACAGGG 466
|||
QY 365 TATCAGGGTTGCTACTATTTCGATCAAGATGGTGAATTTTGTCTGTGAAGGAGAGAGTCCT 424
|||
Db 467 TACAGGGTTGCTATTATTTTTGGTAAAGATGGAAGTTTGTATGTGAAGGAGAGATGAT 526
|||
QY 425 GAACCCAAAGACCACTGCTTA 444
|||
Db 527 GAACCCAAAGGCAAAATATGTA 546
|||

RESULT 10
AY007240 684 bp DNA linear PLN 02-JAN-2001
LOCUS Lycopersicon esculentum proteinase inhibitor II gene, complete cds.
DEFINITION
ACCESSION AY007240
VERSION AY007240.1 GI:12007535
KEYWORDS
SOURCE Lycopersicon esculentum (tomato)
ORGANISM
REFERENCE
1 Xie,X. and Wu,N.
Isolation of tomato proteinase inhibitor II gene and analysis of its structure
Unpublished
2 (bases 1 to 684)
Xie,X. and Wu,N.
Direct Submission
Submitted (17-AUG-2000) Plant Developmental Molecular Biology, Institute of Developmental Biology, Chinese Academy of Sciences, Nanyitiao No.3, Zhongguancun, Haidian District, Beijing 100080, China
FEATURES
source Location/Qualifiers
1..684
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/db_xref="taxon:4081"
join(<9..60,170..564)
/product="proteinase inhibitor II"
join(9..60,170..564)
/codon_start=1
/product="proteinase inhibitor II"
/protein_id="NAG12170.1"
/db_xref="GI:12007536"
/translation="MAVHKVNFVAYLLIVLGMFLYVDKAKTREGCNLGFICPRSE
GSPLPNICNCSGYKCNYNYSFGKFIGESDPRKPNACTFNCDPNIAYSRCPRSQ
GKSLIYPTGCTTCTCTGYKGYFGDKGKFCVCESEDEPKANMYPVM"
ORIGIN
Query Match 49.0%; Score 259.2; DB 8; Length 684;
Best Local Similarity 74.3%; Pred. No. 1.9e-58;
```

```
Matches 327; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 5 ATGGCTGTTCCAAAGTACCTTCCTGCTGCTACTGCTTCTTCTGATGAGTGTCTTA 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 112 ATGGTTTATACATACACAGTAGTGTATATATTTTCTTATATATATTTGTAGGA 171
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 65 CTTCGCAACATCTTGCATGCCAAGGCTTGCTAGAGATGCTGCTATTTAGCTATGCC 124
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 172 ATGTTTCTATATGTTGATGCCAAGGCTTGCTAGAGATGCTGCTATTTAGCTATGCC 231
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 125 ATATGCCCACTGTCAGAGGAAGTCCCAAAACCTATATGACCAAAATTTGCTCAGGC 184
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 232 ATATGCCCACTGTCAGAGGAAGTCCGCTAAATCCCATATGATCAATTTGCTCAGGC 291
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 185 TATAAGGTTGCAACTATTACAGTGTAAAGAGATTTGATTTGTGAAGAGAAATCTGAC 244
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 292 TATAAGGTTGTAATATTATTAATTTCTTCGGAATAATTTATTTGTGAAGAGAAATCTGAT 351
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 245 CCTAGAAACCCAAAGATGTACCTCGAATGATACACAGATTCCTATTCAAAATGT 304
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 352 CCAAAAGGCCAAATGCTTGCACTTTAATTTGTGATCCAAATATTCCTATTCAAGATGT 411
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 305 CTCTGTTTCAGAGGAAGATGATAATTAACCCACTGGATGCAACCACTTGTTCACGGGC 364
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 412 CCGCGTTCAGAGGAATCGTTAATTTATCCACCGGATGATACCACTGTTGCAAGGG 471
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 365 TATCAGGTTGCTACTATTTCGATCAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 424
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 472 TACAAGGTTGCTATTATTATTGTTAAAGATGGAAGTTTGTATGTGAAGGAGAGATGAT 531
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 425 GAACCAAGCACTGCTTA 444
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 532 GAACCAAGCAAAATATGTA 551
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11.
AY129402
LOCUS AY129402 1670 bp DNA linear PLN 16-JUN-2004
DEFINITION Lycopersicon esculentum proteinase inhibitor II gene, complete cds.
ACCESSION AY129402
VERSION AY129402.1 GI:33413549
KEYWORDS Lycopersicon esculentum (tomato)
SOURCE Lycopersicon esculentum
ORGANISM Lycopersicon esculentum
REFERENCE 1 (bases 1 to 1670)
AUTHORS Zhang,H.Y., Xie,X.Z., Xu,Y.Z. and Wu,N.H.
TITLE Isolation and functional assessment of a tomato proteinase inhibitor II gene
JOURNAL Plant Physiol. Biochem. 42 (5), 437-444 (2004)
PUBMED 15191748
REFERENCE 2 (bases 1 to 1670)
AUTHORS Zhang,H., Xie,X., Wu,N. and Huang,M.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-2002) Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Nanyitiao No.3 Zhongguancun, Beijing 2707, China
FEATURES
    source
        1..1670
            /organism="Lycopersicon esculentum"
            /mol_type="genomic DNA"
            /db_xref="taxon:4081"
            join(-995..1046,1156..>1550)
            /product="proteinase inhibitor II"
            join(995..1046,1156..1550)
            /codon_start=1
            /product="proteinase inhibitor II"
            /protein_id="AA05016.1"
            /db_xref="GI:33413550"
    mRNA
    CDS
        translations="MAVHKVNFVAYLLVLGMFLYVDKACTRECNGLFGICPRSE
        GSPNLPICNCCSGYKGCYNYNSFGKFCIEGESDPRKPNACTFNCDPNIAYSRCPRSE
```

```
ORIGIN
Query Match 49.0%; Score 259.2; DB 8; Length 1670;
Best Local Similarity 74.3%; Pred. No. 1.8e-58;
Matches 327; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 5 ATGGCTGTTCCAAAGTACCTTCCTGCTGCTACTGCTTCTTCTGATGAGTGTCTTA 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1098 ATGGTTTATACATACACAGTAGTGTATATATTTTCTTATATATATTTGTAGGA 1157
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 65 CTTCGCAACATCTTGCATGCCAAGGCTTGCTAGAGATGCTGCTATTTAGCTATGCC 124
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1158 ATGTTTCTATATGTTGATGCCAAGGCTTGCTAGAGATGCTGCTATTTAGCTATGCC 1217
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 125 ATATGCCCACTGTCAGAGGAAGTCCCAAAACCTATATGACCAAAATTTGCTCAGGC 184
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1218 ATATGCCCACTGTCAGAGGAAGTCCGCTAAATCCCATATGCAATCAATTTGCTCAGGC 1277
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 185 TATAAGGTTGCAACTATTACAGTGTAAAGAGATTTGATTTGTGAAGAGAAATCTGAC 244
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1278 TATAAGGTTGTAATATTATTAATTTCTTCGGAATAATTTATTTGTGAAGAGAAATCTGAT 1337
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 245 CTCTAGAAACCCAAAGATGTACCTTCGAATGATACACAGATTCCTATTCAAAATGT 304
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1338 CCAAAAGGCCAAATGCTTGCACTTTAATTTGTGATCCAAATATTCCTATTCAAGATGT 1397
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 305 CTCTGTTTCAGAGGAAGATGATAATTAACCCACTGGATGCAACCACTTGTTCACGGGC 364
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1398 CCGCGTTCAGAGGAATCGTTAATTTATCCACCGGATGATACCACTGTTGCAAGGG 1457
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 365 TATCAGGTTGCTACTATTTCGATCAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 424
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1458 TACAAGGTTGCTATTATTATTGTTAAAGATGGAAGTTTGTATGTGAAGGAGAGATGAT 1517
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 425 GAACCAAGCACTGCTTA 444
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1518 GAACCAAGCAAAATATGTA 1537
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12.
POTI1KA
LOCUS POTI1KA 1241 bp DNA linear PLN 28-APR-1994
DEFINITION S.tuberosum (potato) inhibitor II-chloramphenicol acetyltransferase gene (IIK), complete cds.
ACCESSION M15186
VERSION M15186.1 GI:169488
KEYWORDS inhibitor II-chloramphenicol acetyl transferase.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
REFERENCE 1 (bases 1 to 1241)
AUTHORS Thornburg,R.W., An,G., Cleveland,T.E., Johnson,R. and Ryan,C.A.
TITLE Wound-inducible expression of a potato inhibitor II-chloramphenicol acetyltransferase gene fusion in transgenic tobacco plants
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 84, 744-748 (1987)
COMMENT Original source text: Solanum tuberosum DNA.
FEATURES
    source
        1..1241
            /organism="Solanum tuberosum"
            /mol_type="genomic DNA"
            /db_xref="taxon:4113"
            prim_transcript 258..1042
            exon
                /note="inhibitor IIK mRNA and introns"
                258..357
                /gene="IIK"
                /number=1
                join(306..357,475..884)
                /gene="IIK"
                /note="precursor"
                /codon_start=1
    CDS
```

/product="inhibitor II-chloramphenicol acetyltransferase"
 /protein_id="AAA33815.1"
 /db_xref="GI:169489"
 /translation="MDVHKEVNFVAYLLVLGLLVLSAMDVDAKACIRECGNLGPGI
 CPSESPENPITCNCCAGYKGCNYSANGAFICGSDPKPKAPLNCDDPHAYSK
 CPSESKSLIYPTGCTTCCTGCGYKGCYFNGKGFVCEGESDEPKANNYPAM"
 join(306..357,475..512)
 /gene="IIK"
 /size=881
 /gene="IIK"
 /product="inhibitor II-chloramphenicol acetyltransferase"
 join(306..357,475..512)
 /gene="IIK"
 358..474
 /gene="IIK"
 /number=1
 475..1042
 /gene="IIK"

ORIGIN

Query Match 48.8%; Score 258; DB 8; Length 1241;
 Best Local Similarity 74.7%; Pred. No. 3.8e-58;
 Matches 324; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

Qy 11 GTTCACAAAGTTAGCTTCCTTGCTGCTACTTGTCTTGGATGGATGTTTCTACTTGGC 70
 Db 438 GTTTATATTTTCCCTTTATATATATTTGTTGTAGGATTTATGTTGAAGCGC 497
 Qy 71 AAACATGTTGATGCCAAGGCTTGCTAGAGAAATGGTGCATTTTAGCTATGGCATATGC 130
 Db 498 ATGGATGTTGATGCCAAGGCTTGCTAGAGAAATGGTGCATTTTAGCTATGGCATATGC 557
 Qy 131 CCAGTTTCAGAGGAAGTCCCAAAACCTATATGACCAATTTGCTCAGGCTATAAG 190
 Db 558 CCAGTTTCAGAGGAAGTCCGAAATCCGATATGCCACCACTGTTGTGCAAGTTATAA 617
 Qy 191 GGTTCACAACTATTACAGTGTCTAAAGGAGATTGATTTGTGAAGGAGAAATCTGACCCCTAGA 250
 Db 618 GGTTCACAACTATTATAGTGCAAATGGGCTTTCATTTGTGAAGGACAACTCTGACCCCAAAA 677
 Qy 251 AACCCAAAAGATTGTACCTTGAATGATGATACAGAGTTGCTTATTTCAAAATGCTCCTCGT 310
 Db 678 AACCCAAAAGCATGCCCTTAAATTTGGATGCCATATTTGCCACTCAAAAGTGTCCCCGT 737
 Qy 311 TCAGAGGAAGATGATAATTAACCCACCTGGATGACCACTTGTTCGACGGGCTATCAG 370
 Db 738 TCAGAGGAAGATGCTTAATTTATCCACCAGATGACCACTGCTGCACAGGATCAAG 797
 Qy 371 GGTTCCTACTATTTCGATCAAGATGGTGTATTTGTCTGTGAAGGAGAGAGTCTCTGAACCC 430
 Db 798 GGTTCCTACTATTTCGTAATAATGGCAAGTTTGTATGTGAAGGAGAGATGATGAGCCC 857
 Qy 431 AAGACCACTGCTTA 444
 Db 858 AAGCAAAATATGTA 871

RESULT 13

AY247794
 LOCUS Solanum phureja 482 bp mRNA linear PLN 09-APR-2003
 DEFINITION Solanum phureja proteinase inhibitor II protein (PIN-112x) mRNA,
 complete cds.
 ACCESSION AY247794
 VERSION AY247794.1 GI:29650886
 KEYWORDS Solanum phureja
 SOURCE Solanum phureja
 ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 482)
 Bu, Q., Yang, S. and Wan, J.
 Potato proteinase inhibitor mRNA, complete cds

JOURNAL REFERENCE

2 (bases 1 to 482)
 Bu, Q., Yang, S. and Wan, J.
 Direct Submission
 Submitted (03-MAR-2003) National Key Laboratory of Crop Genetics
 and Germplasm Enhancement, Rice Research Institute, Weigang,
 Nanjing, Jiangsu 210095, China
 Location/Qualifiers

FEATURES

source

1..482
 /organism="Solanum phureja"
 /mol_type="mRNA"
 /variety="IVP101"
 /db_xref="taxon:172790"
 /tissue type="wound-induced leaf"
 /note="diploid potato cultivar"
 1..482
 /gene="PIN-112x"
 3..467
 /gene="PIN-112x"
 /codon_start=1
 /product="proteinase inhibitor II protein"
 /protein_id="AB088244.1"
 /db_xref="GI:29650887"
 /translation="MAVHKEVNFVAYLLVLGLLVLSAMEHVDKACTLECGNLGFG
 ICPSESPENRITCNCCAGYKGCNYSANGAFICGSDPKPNPVCPRNCDTNIAYS
 KCPSESKSLIYPTGCTTCCTGCGYKGCYFNGKGFVCEGESDEPKANNYPAM"

gene

CDS

ORIGIN

Query Match 48.7%; Score 257.6; DB 8; Length 482;
 Best Local Similarity 76.4%; Pred. No. 5e-58;
 Matches 333; Conservative 0; Mismatches 94; Indels 9; Gaps 1;

Qy 18 AAGTTAGCTTCCTTGCTGCTACTTGTCTTGGATGGATGTTTCTACTTTG----- 68
 Db 19 AAGTTAATTTGTTGCTTACTACTTAATTTGTTCTTGGATTTTGGTACTTTGAAGCGCA 78
 Qy 69 CGAAACATGTTGATGCCAAGGCTTGCTAGAGAAATGGTTCATTTTAGCTATGGCATAT 128
 Db 79 TGGAGCATGTTGATGCCAAGGCTTGCACTTTAGAAATGTTGATTTGGGTTTGGGATAT 138
 Qy 129 GCCACGTTTCAGAGGAAGTCCCAAAACCTATATGCAACCAATTTTGTCTGAGGTATA 198
 Db 139 GCCACGTTTCAGAGGAAGTCCCGAAATTCGCATATGCAACCACTGTTGTGCAAGTTATA 198
 Qy 189 AGGTTGCAACTATTACAGTGTCTAAAGGAGATTGATTTGTGAAGGAGAAATCTGACCCCTA 248
 Db 199 AAGTTGCAATTTATATAGTGCAAATGGGCTTTCAATTTGCGAAGGAGATCTGACCCCA 258
 Qy 249 GAAACCCAAAAGATTGCTTTCGAAATGTATACACAGATTGCTTATTCAAAATGTCCTC 308
 Db 259 AAAACCCAAAATGTTTCCCGCCGAAATTTGTATACAAATATTTGCTATTCAAAGTGTCCCC 318
 Qy 309 GTTCAGAGGAAGATGATAATTAACCCACCTGGATGCAACCACTTTTGCACCGGCTATC 368
 Db 319 GTTCAGAGGAAGAAATCGCTAAATTTATCCACCGGATGTACCACTGCTGCACAGGTTACA 378
 Qy 369 AGGTTGCTACTATTTCGATCAAGATGCTGATTTGCTGTGAAGGAGAGAGTCTCTGAAC 428
 Db 379 AGGTTGCTACTATTTCGGTAAATGCGCAAGTTTGTATGTGAAGGAGAGATGATGAGC 438
 Qy 429 CCAAGACCACTGCTTA 444
 Db 439 CCAAGCAAAATATGTA 454

RESULT 14

STPIIR1
 LOCUS Solanum phureja 512 bp mRNA linear PLN 12-SEP-1993
 DEFINITION Potato (Solanum tuberosum) mRNA 1 for proteinase inhibitor II.
 ACCESSION X03778
 VERSION X03778.1 GI:21523
 KEYWORDS protease inhibitor,
 Solanum tuberosum (potato)

ORGANISM Solanum tuberosum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum.
AUTHORS Sanchez-Serrano, J., Schmidt, R., Schell, J. and Willmitzer, L.
TITLE Nucleotide sequence of proteinase inhibitor II encoding cDNA of potato (Solanum tuberosum) and its mode of expression
JOURNAL Mol. Gen. Genet. 203, 15-20 (1986)
COMMENT Data kindly reviewed (15-JUN-1986) by J. Sanchez-Serrano.
FEATURES Location/Qualifiers
 source 1..512
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /db_xref="taxon:4113"
 CDS 48..512
 /notes="unnamed protein product; put. proteinase inhibitor II (aa 1-154)"
 /codon_start=1
 /protein_id="CAA27408.1"
 /db_xref="GI:21524"
 /db_xref="GOA:P01080"
 /translaton="MDVHKRVNFAYLLVLGLLVLSAMEHVDKACTLECGNLGFGICPRSGSPENRICNCAGYGCNYSANGAFICGSQDPKKPKACPLNCDPHIAYS KCPRSEGSILYPTGTCTTCTGYKGKGFVCEGESDEPKANMYPAM"

	Query Match	48.5%	Score 256.8;	DB 8;	Length 512;
	Best Local Similarity	76.1%;	Pred. No. 8.2e-58;		
	Matches 332;	Conservative 0;	Mismatches 95;	Indels 9;	Gaps 1;

```

Qy      18 AAGTTAGTCCTTCCTTGCTGCTACTCTTCTCTGGATGGATGTTTCTACTTGG----- 68
Db      '64 AGTTAAITTCGTTGCTTACTCTACTAATTGTTCTTGGATTATTTGGTACTTTGTAAGCCGA 123

Qy      69 CGAAACATGTGTGATGCCAAGCTTGTACTAGAGAATGTGTGCAATTTTAGCTATGGCATAT 128
Db      124 TGGAGCATGTTGATGCCAAGGCTTGCACTTTAGAAATGTGGTAATCTTTGGGTTTTGGGATAT 183

Qy      129 GCCACGTTTCAGBAGGAAGTCCCCAAAACCTATATGCACCATTGTTGCTCAGGCTATA 188
Db      184 GCCACGTTTCAGBAGGAAGTCCGAAAATCGCATATGCACCACTGTTGTGCAGGTTATA 243

Qy      189 AGGGTTCCAACTATTACAGTGCTAAAGGAGATTGATTTGTGAAGGAGAATCTGCACCTA 248
Db      244 AAGTTGCAATTATTATAGTGCAATGGGCTTTCATTTGTGAAGGACATCTGACCCA 303

Qy      249 GAAACCCAAAAGATTTGTACCTTCGAATGTGATACACAGATTGCTTATTCAAAATGTCCTC 308
Db      304 AAAAACCCAAAAGCATGCCCTTAAATTTGCGATCCACATATTGCTACTCAAAGTGTCCCC 363

Qy      309 GTTCAGBAGGAAGATGATTAATTAACCCACTGGATGCACACTTGTTCACCGGCTATC 368
Db      364 GTTCAGBAGGAANAATCGCTAAATTTATCCCACCGGATGTACCAATGCTGCACAGGGTACA 423

Qy      369 AGGGTTCCTACTATTTCGATCAAGATGGTGATTTTGTCTGTAAGGAGAGAGTCTTGAAC 428
Db      424 AGGGTTCCTACTATTTCGTTAAAAATGCGAAGTTTGTATGTGAAGGAGAGATGATGAGC 483

Qy      429 CCAAGACCACTGCTTA 444
Db      484 CCAAGGCAATATGTA 499

```

RESULT 15
POTPPINH POTPPINH 666 bp mRNA linear PLN 28-NOV-1994
LOCUS Solanum tuberosum proteinase inhibitor II mRNA, complete cds.
DEFINITION L37519
ACCESSION L37519
VERSION L37519.1 GI:576528
KEYWORDS double-headed proteinase inhibitor; proteinase inhibitor II.
SOURCE Solanum tuberosum (potato)

Search completed: January 22, 2005, 06:33:37
Job time : 2766 secs

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 04:42:11 ; Search time 2581 Seconds
(without alignments)
7468.663 Million cell updates/sec

Title: US-10-725-829-1
Perfect score: 529
Sequence: 1 cataatggctgttcacaag.....ctgtaatgtggactttattg 529

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gse1.*
9: gb_gse2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	362.8	68.6	545	4	BI176253
2	362.8	68.6	554	2	BF053417
3	362.8	68.6	571	5	BQ114991
4	337.4	63.8	451	5	BQ114990
5	289.6	54.7	469	2	AW035333
6	289.6	54.7	614	4	BI421156
7	289.6	54.7	626	4	BI205417
8	289.6	54.7	663	4	BI205603
9	270.8	51.2	625	4	BI433294
10	270.8	51.2	628	4	BI433271
11	269.4	50.9	672	2	BF133853
12	269.4	50.9	675	4	BI436259
13	269.4	50.9	677	2	BF153879
14	267.8	50.6	591	2	BF153508
15	267.2	50.5	779	4	BI421162
16	267	50.5	641	2	BF460374
17	267	50.5	655	4	BI406670
18	265.4	50.2	477	2	AW093009
19	265.4	50.2	484	2	BE463283
20	265.4	50.2	501	1	AI771251
21	265.4	50.2	503	2	AW221025
22	265.4	50.2	504	1	AI485972
23	265.4	50.2	515	1	AI489819
24	265.4	50.2	533	1	AI771306

25	265.4	50.2	550	1	AI485979	EST244300
26	265.4	50.2	558	2	AW092640	AW092640
27	265.4	50.2	564	2	BE354745	BE354745
28	265.4	50.2	571	1	AI771944	AI771944
29	265.4	50.2	581	1	BI929125	BI929125
30	265.4	50.2	582	2	BE354912	BE354912
31	265.4	50.2	587	1	AI484805	AI484805
32	265.4	50.2	604	1	AI487479	AI487479
33	265.4	50.2	612	2	BF460196	BF460196
34	265.4	50.2	623	4	BI930668	BI930668
35	265.4	50.2	623	4	BI930919	BI930919
36	265.4	50.2	623	4	BI932864	BI932864
37	265.4	50.2	624	2	AW093425	AW093425
38	265.4	50.2	626	2	BI929471	BI929471
39	265.4	50.2	629	2	AW217790	AW217790
40	265.4	50.2	630	4	BI931013	BI931013
41	265.4	50.2	637	2	BE463161	BE463161
42	265.4	50.2	644	4	BI935830	BI935830
43	265.4	50.2	645	4	BI935289	BI935289
44	265.4	50.2	655	4	BI930423	BI930423
45	265.4	50.2	655	4	BI933036	BI933036

ALIGNMENTS

RESULT 1
BI176253
LOCUS EST517336 cSTS Solanum tuberosum cDNA clone cSTS28K7 5' sequence, linear EST 07-MAR-2003
DEFINITION mRNA sequence.
ACCESSION BI176253
VERSION BI176253.1 GI:14642064
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 545)
AUTHORS van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Chieming, A., Bougri, O., Buell, C.R., Ronning, C., Tanksley, S. and Baker, B.
TITLE Generations of ESTs from sprouting potato eyes
JOURNAL Unpublished (2000)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13P-R.

FEATURES

source Location/Qualifiers
1..545
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS28K7"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/clone_lib="cSTS"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."

ORIGIN

Query Match 68.6%; Score 362.8; DB 4; Length 545;
Best Local Similarity 84.6%; Pred. No. 4.4e-89;

Matches 445; Conservative 0; Mismatches 72; Indels 9; Gaps 3;

```
QY 1 CATAATGGCTGTTACAAAGTCTAGCTTCCTTGCCTACTTGTCTTGGATGGATGTT 60
    |||||
Db 6 CATAATGGCTGTTACAAAGTCTAGCTTCCTTGCCTACTTGTCTTGGATGGATGTT 62
    |||||
QY 61 TCTACTTGGCAACATGTTGATGCCAAGGCTTCTACTAGAGAATGTTGTCATTTAGCTA 120
    |||||
Db 63 TCTTCTTGTAAACACTGTTTGAAGGAGTGTCTCAAGAAATATGTTGTAATTTGGGTA 122
    |||||
QY 121 TGGCATATGCCACGTTTCAGAGGAAGTCCCAAAAACCTATATGACCAATTTGTGCTC 180
    |||||
Db 123 TGGCATATGCCACGTTTCAGAGGAAGTCCCAAAAACCTATATGACCAATTTGTGCTC 182
    |||||
QY 181 AGGCTATAAGGTTGCAACTATTACAGTGTCTAAAGGAGATTTGATTTGTGAAGGAGATC 240
    |||||
Db 183 GGGCTACAGGGTTGCAAAATATTACAGTGTCTAAAGGAGATTTGATTTGTGAAGGAGATC 242
    |||||
QY 241 TGACCCCTAGAAAACCCAAAAGATTGTACCTTCGAATGTGATACACAGATTCCTTATTTCAA 300
    |||||
Db 243 TGACCCCTAGAAAACCCAAAAGATTGTACCTTCGAATGTGATACACAAATTCCTTATTTCAA 302
    |||||
QY 301 ATGTCCTCGTTTCAGAGGAAGATGATATTAACCACTGGATGACCACTTTGTGCTC 360
    |||||
Db 303 ATGTCCTCGTTTCAGAGGAAGATGATATTAACCACTGGATGACCACTTTGTGCTC 362
    |||||
QY 361 GGGCTATCAGGTTGCTACTATTTCGATCAAGATGTGATTTTGTCTGAAGGAGAG 420
    |||||
Db 363 CGGCTACAGGGTTGCTACTATTTCGATCAAGAGTGTGATTTTGTCTGAAGGAGAG 422
    |||||
QY 421 TCGTGAACCCAAAGACCACTGCT----TATTCTTAATCAATCATATGTTTATCTATCAA 476
    |||||
Db 423 TCGTGAACCCAAAGCGAGTTGCTTAAAGTATTCTTAATTAATCGTTTGTGCAATTTATTAG 482
    |||||
QY 477 AAAAAAATATGATGATGATATATGCTGTTACTGTTACTGTAATGCTGCAC 522
    |||||
Db 483 TAACAA--ATGTGTGATTATATATGCTTAATCTTTGTGAAGGAGAG 526
    |||||
```

RESULT 2
BF053417
LOCUS
DEFINITION
EST343647 potato leaves and petioles Solanum tuberosum cDNA clone
cSTB35N6 5' sequence, mRNA sequence.

ACCESSION
BF053417

VERSION
BF053417.1 GI:10807313

KEYWORDS
EST.

SOURCE
Solanum tuberosum (potato)

ORGANISM
Solanum tuberosum

REFERENCE
AUTHORS
van der Hoeven, R.S., Bezzerides, J., Holt, I.E., Liang, F., Cho, J.,
Utterback, T., Hansen, B., Doan, B., Bougri, O., Buehl, C.R.,
Roning, C.M., Fry, W.E., Tanksley, S.D. and Baker, B.

TITLE
JOURNAL
COMMENT
Generation of ESTs from potato leaves and petioles
Unpublished (2000)

Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:

http://genome.arizona.edu/orders/
Location/Qualifiers

FEATURES
source
1. .554

/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"

/db_xref="taxon:4113"
/clone="cSTB35N6"
/tissue_type="leaflets and petioles"

/dev_stage="8 weeks old plants"

/lab_host="SOLR"

/clone_lib="potato leaves and petioles"

/note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."

Query Match 68.6%; Score 362.8; DB 2; Length 554;
Best Local Similarity 84.6%; Pred. No. 4.4e-99;
Matches 445; Conservative 0; Mismatches 72; Indels 9; Gaps 3;

ORIGIN

```
QY 1 CATAATGGCTGTTACAAAGTCTAGCTTCCTTGCCTACTTGTCTTGGATGGATGTT 60
    |||||
Db 6 CATAATGGCTGTTACAAAGTCTAGCTTCCTTGCCTACTTGTCTTGGATGGATGTT 62
    |||||
QY 61 TCTACTTGGCAACATGTTGATGCCAAGGCTTCTACTAGAGAATGTTGTCATTTAGCTA 120
    |||||
Db 63 TCTTCTTGTAAACACTGTTTGAAGGAGTGTCTCAAGAAATATGTTGTAATTTGGGTA 122
    |||||
QY 121 TGGCATATGCCACGTTTCAGAGGAAGTCCCAAAAACCTATATGACCAATTTGTGCTC 180
    |||||
Db 123 TGGCATATGCCACGTTTCAGAGGAAGTCCCAAAAACCTATATGACCAATTTGTGCTC 182
    |||||
QY 181 AGGCTATAAGGTTGCAACTATTACAGTGTCTAAAGGAGATTTGATTTGTGAAGGAGATC 240
    |||||
Db 183 GGGCTACAGGGTTGCAAAATATTACAGTGTCTAAAGGAGATTTGATTTGTGAAGGAGATC 242
    |||||
QY 241 TGACCCCTAGAAAACCCAAAAGATTGTACCTTCGAATGTGATACACAGATTCCTTATTTCAA 300
    |||||
Db 243 TGACCCCTAGAAAACCCAAAAGATTGTACCTTCGAATGTGATACACAAATTCCTTATTTCAA 302
    |||||
QY 301 ATGTCCTCGTTTCAGAGGAAGATGATATTAACCACTGGATGACCACTTTGTGCTC 360
    |||||
Db 303 ATGTCCTCGTTTCAGAGGAAGATGATATTAACCACTGGATGACCACTTTGTGCTC 362
    |||||
QY 361 GGGCTATCAGGTTGCTACTATTTCGATCAAGATGTGATTTTGTCTGAAGGAGAG 420
    |||||
Db 363 CGGCTACAGGGTTGCTACTATTTCGATCAAGAGTGTGATTTTGTCTGAAGGAGAG 422
    |||||
QY 421 TCGTGAACCCAAAGACCACTGCT----TATTCTTAATCAATCATATGTTTATCTATCAA 476
    |||||
Db 423 TCGTGAACCCAAAGCGAGTTGCTTAAAGTATTCTTAATTAATCGTTTGTGCAATTTATTAG 482
    |||||
QY 477 AAAAAAATATGATGATGATATATGCTGTTACTGTTACTGTAATGCTGCAC 522
    |||||
Db 483 TAACAA--ATGTGTGATTATATATGCTTAATCTTTGTGAAGGAGAG 526
    |||||
```

RESULT 3

BQ114991/c

LOCUS

DEFINITION

EST600567 mixed potato tissues Solanum tuberosum cDNA clone STMCX74

3' end, mRNA sequence.

ACCESSION

BQ114991

VERSION

BQ114991.1 GI:20166953

KEYWORDS

SOURCE

EST.

ORGANISM

Solanum tuberosum (potato)

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamids; Solanales; Solanaceae; Solanum.

1 (bases 1 to 571)

Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C.,

Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and

Karamycheva, S.A.

Generation of a set of potato cDNA clones for microarray analyses

Unpublished (2002)

Contact: Robin Buell

The Institute for Genomic Research


```

9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T7.

FEATURES
    source
        1..571
            /organism="Solanum tuberosum"
            /mol_type="mRNA"
            /cultivar="Kennebec or Binjete"
            /db_xref="taxon:4113"
            /clone="STMCK74"
            /tissue_type="mixed tissues"
            /lab_host="SOLR"
            /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
            XhoI; supplier: Combination of untreated and Phytophthora
            infestans-treated libraries of stolons, leaves, leaflets,
            axillary buds of stem explants, petioles, germinating
            eyes, tubers, or roots."

ORIGIN
    Query Match      68.6%; Score 362.8; DB 5; Length 571;
    Best Local Similarity 84.6%; Pred. No. 4.5e-89;
    Matches 445; Conservative 0; Mismatches 72; Indels 9; Gaps 3;

QY 1 CATAATGGCTGTTCACAAAGTTAGCTTCCTTGCTGCTACTTGTCTTGGATGGATGTT 60
Db 552 CATAATGGCTGTTCACAAAGTTAGCTTCCTTGCTGCTACTTGTCTTGGATGTT 496

QY 61 TCTACTTGGCAACATGTTGATGCCAAGGTTCTACTAGAGATGGTCATTTAGCTA 120
Db 495 TCTTCTTGAACACTGTTGATGCCAAGGTTCTCAAGAATATGTGTAATTTGGGTA 436

QY 121 TGGCATATGCCACGTTCTGAAGGAGTCCCAAAACCTATATGACCAATTTGTGCTC 180
Db 435 TGGACTATGCCACGTTCTGAAGGAGTCCGHAAGGCCCATATGACCAATTTGTGCTC 376

QY 181 AGGCTATAAGGGTTGCAACTATTACAGTGTCTAAAGGAGATTTGATTTGTGAAGGAGATC 240
Db 375 GGCTACAGGGTTGCANATATTACAGTGTCTACGAGATTTAGTTTGTGAAGGGAGTC 316

QY 241 TGACCTTAGAAACCCAAAGATTTGATCTTGAATGTGATACACAGATGCTTATTCAAA 300
Db 315 TGACCCAGAAACCAAAATGATTTCTTATGATGTGATACACAAATTCCTTATTCAAA 256

QY 301 ATGCTCTCGTTCAGAGGAAGATGATNATTAACCCACTGGATGCCACTTGTTCAC 360
Db 255 ATGCTCTCGTTCAGAGGAAGATGATNATTAACCCACTGGATGCCACTTGTTCAC 196

QY 361 GGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGTGATTTGTCTGAAGGAGAG 420
Db 195 CGGCTACAGGGTTGCTACTATTTCGATCAAGAGGTGATTTGTGTGAAGGAGAG 136

QY 421 TCCCTGAACCAAGACCACTGCT----TATTCTTAATCAATCATATGTTGTTATCA 476
Db 135 TCCCTGAACCAAGCAGTGTCTTAAGTATTCTTAATTAATCGTTTGTGCAATTTATTAG 76

QY 477 AAAAAAATATGATGATGATGATATATCTGTTGTTACTGTTATATGTTGAC 522
Db 75 TAACAA--ATTGTGTGATTATATATGCTAATCTTTGTGAAGGAGAG 32

RESULT 4
BQ114990
LOCUS EST600566 mixed potato tissues Solanum tuberosum cDNA clone STMCK74
DEFINITION 5' end, mRNA sequence.
ACCESSION BQ114990
VERSION BQ114990.2 GI:21916736
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)

ORGANISM
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
REFERENCE
1 (bases 1 to 451)
AUTHORS
Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C.,
Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and
Karamycheva, S.A.
Generation of a set of potato cDNA clones for microarray analyses
Unpublished (2002)
On Apr 17, 2002 this sequence version replaced gi:20166952.
Contact: Robin Buell
The Institute for Genomic Research
712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T7.
Location/Qualifiers
    source
        1..451
            /organism="Solanum tuberosum"
            /mol_type="mRNA"
            /cultivar="Kennebec or Binjete"
            /db_xref="taxon:4113"
            /clone="STMCK74"
            /tissue_type="mixed tissues"
            /lab_host="SOLR"
            /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
            XhoI; supplier: Combination of untreated and Phytophthora
            infestans-treated libraries of stolons, leaves, leaflets,
            axillary buds of stem explants, petioles, germinating
            eyes, tubers, or roots."

ORIGIN
    Query Match      63.8%; Score 337.4; DB 5; Length 451;
    Best Local Similarity 87.6%; Pred. No. 4.3e-82;
    Matches 381; Conservative 0; Mismatches 51; Indels 3; Gaps 1;

QY 1 CATAATGGCTGTTCACAAAGTTAGCTTCCTTGCTGCTACTTGTCTTGGATGGATGTT 60
Db 20 CATAATGGCTGTTCACAAAGTTAGCTTCCTTGCTGCTACTTGTCTTGGATGTT 76

QY 61 TCTACTTGGCAACATGTTGATGCCAAGGTTCTACTAGAGATTTGTCATTTAGCTA 120
Db 77 TCTTCTTGAACACTGTTGATGCCAAGGTTCTCAAGAATATGTGTAATTTGGGTA 136

QY 121 TGGCATATGCCACGTTCTGAAGGAGTCCCAAAACCTATATGACCAATTTGTGCTC 180
Db 137 TGGACTATGCCACGTTCTGAAGGAGTCCGHAAGGCCCATATGACCAATTTGTGCTC 196

QY 181 AGGCTATAAGGGTTGCAACTATTACAGTGTCTAAAGGAGATTTGATTTGTGAAGGAGATC 240
Db 197 GGCTACAGGGTTGCTACTATTTCGATCAAGATGGTGTGATTTGTCTGAAGGAGAG 256

QY 241 TGACCTTAGAAACCCAAAGATTTGATCTTGAATGTGATACACAGATGCTTATTCAAA 300
Db 257 TGACCCAGAAACCAAAATGATTTCTTATGATGTGATACACAAATTCCTTATTCAAA 316

QY 301 ATGCTCTCGTTCAGAGGAAGATGATNATTAACCCACTGGATGCCACTTGTTCAC 360
Db 317 ATGCTCTCGTTCAGAGGAAGATGATNATTAACCCACTGGATGCCACTTGTTCAC 376

QY 361 GGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGTGATTTGTCTGAAGGAGAG 420
Db 377 CGGCTACAGGGTTGCTACTATTTCGATCAAGAGGTGATTTGTGTGAAGGAGAG 436

QY 421 TCCCTGAACCAAGAC 435
Db 437 TCCCTGAACCAAGGC 451

```

```

RESULT 5
AW035333
LOCUS
DEFINITION
  EST280696 tomato callus, TAMU Lycopersicon esculentum cDNA clone
  cLEC40H22 similar to protein inhibitor II, mRNA sequence.
ACCESSION
AW035333
VERSION
AW035333.1 GI:5894089
SOURCE
  Lycopersicon esculentum (tomato)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
  1 (bases 1 to 469)
REFERENCE
  Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
  Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S.,
  Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
  Giovannoni,J.
  Generation of ESTs from tomato callus tissue
  Unpublished (1999)
  Contact: CUGI
  Clemson University
  100 Jordan Hall, Clemson, SC 29634, USA
  Email: http://www.genome.clemson.edu/orders/index.html
  5 prime sequence.
FEATURES
  source
    Location/Qualifiers
      1..469
        /organism="Lycopersicon esculentum"
        /mol_type="mRNA"
        /cultivar="TA496"
        /db_xref="taxon:4081"
        /clone="cLEC40H22"
        /tissue_type="callus"
        /dev_stage="25-40 days old"
        /lab_host="XLI-Blue MRF"
        /clone_lib="tomato callus, TAMU"
        /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
        XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
        of seedlings 7-10 days post-germination were excised, cut
        at both ends and placed on MS medium with no selection.
        Mixed callus was harvested at 25 and 40 days and included
        undifferentiated masses. Tomato Callus EST Library"
ORIGIN
  Query Match      54.7%; Score 289.6; DB 2; Length 469;
  Best Local Similarity 79.4%; Pred. No. 6.9e-69;
  Matches 343; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY  2  ATAAATGGCTGTTTACAAAGTTAGCTTCCTTGCTTGCTACTTCTTGGATGGATGTTT 61
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  32 ACATAGGCTGTTTACAAAGTTAGTTTCTTGCTCAGCTACTTGTCTTGGATGTATCTA 91
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  62 CTACTTGCAGAACATGTTGATGCCAAGGCTTGTAAGAGATGTGGTCAATTTAGCTAT 121
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  92 CTAGTAAGCAGCGTGGACACGCTAATGCTTGTAACCAAGATGTGGTAATCTTGCTAT 151
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  122 GGCATATGCCACGTTTCAGAGGAAGTCCCAAAAACCTATATGCACCAATTTGCTCTCA 181
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  152 GGGATATGCCACGTTTCAGAGGAAGTCCAGAAAATCCAATATGTACCAATTTGCTCT 211
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  182 GGCATATAAGGGTTGCAACTATTACAGTGTCTTAAAGGAGATTGATTGTTGAAGGAGATCT 241
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  212 GGCCTAAGGGTTGCAACTATTATACGCTAATGGAACCTTTTATTGTGAGGAAGCTCT 271
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  242 GACCCCTAGAAACCCAAAGATTGTACCTTCGAATGTGATACACAGAGTTGCTTATCAAAA 301
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  272 GATCCAAAAATCCTAACATTTTGCCTCATATTGTGATGCCAACAAATTTGCTATTCAAAG 331
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  302 TGTCTCGTTTCAGAGGAAGATGATAATTAACCCACCTGGATGCACCACTTGTTCACG 361
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  332 TGTCCACGTTTCAGAGGAAGACGATAATCTATCCACAGGATGTACGACGTTGTGCACT 391
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY  362 GGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGATTTTCTCTGTGAAGAGAGT 421
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  392 GGTTCACAAAGGTTGCTACTATTTTGGTCAAGATGGAGAGTTTGTGTGAAGGAGAGT 451
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  422 CTGTGAACCCCAAG 433
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  452 ATTGAACCTAAG 463
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
BI421156
LOCUS
DEFINITION
  EST31822 tomato callus, TAMU Lycopersicon esculentum cDNA clone
  cLEC66G19 5' end, mRNA sequence.
ACCESSION
BI421156
VERSION
BI421156.1 GI:15194430
KEYWORDS
EST.
SOURCE
  Lycopersicon esculentum (tomato)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
  1 (bases 1 to 614)
REFERENCE
  Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
  Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S.,
  Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
  Giovannoni,J.
  Generation of ESTs from tomato callus tissue
  Unpublished (1999)
  Contact: CUGI
  Clemson University
  100 Jordan Hall, Clemson, SC 29634, USA
  Email: http://www.genome.clemson.edu/orders/index.html.
  Location/Qualifiers
    1..614
      /organism="Lycopersicon esculentum"
      /mol_type="mRNA"
      /cultivar="TA496"
      /db_xref="taxon:4081"
      /clone="cLEC66G19"
      /tissue_type="callus"
      /dev_stage="25-40 days old"
      /lab_host="XLI-Blue MRF"
      /clone_lib="tomato callus, TAMU"
      /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
      XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
      of seedlings 7-10 days post-germination were excised, cut
      at both ends and placed on MS medium with no selection.
      Mixed callus was harvested at 25 and 40 days and included
      undifferentiated masses. Tomato Callus EST Library"
ORIGIN
  Query Match      54.7%; Score 289.6; DB 4; Length 614;
  Best Local Similarity 79.4%; Pred. No. 7.3e-69;
  Matches 343; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY  2  ATAAATGGCTGTTTACAAAGTTAGCTTCCTTGCTTGCTACTTCTTGGATGGATGTTT 61
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  13 ACAATGGCTGTTTACAAAGTTAGTTTCTTGCTCAGCTACTTGTCTTGGATGTATCTA 72
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  62 CTACTTGCAGAACATGTTGATGCCAAGGCTTGTAAGAGATGTGGTCAATTTAGCTAT 121
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  73 CTAGTAAGCAGCGTGGACACGCTAATGCTTGTAACCAAGATGTGGTAATCTTGCTAT 132
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  122 GGCATATGCCACGTTTCAGAGGAAGTCCCAAAAACCTATATGCACCAATTTGCTCA 181
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  133 GGGATATGCCACGTTTCAGAGGAAGTCCAGAAAATCCAATATGTACCAATTTGCTCT 192
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  182 GGCATATAAGGGTTGCAACTATTACAGTGTCTTAAAGGAGATTGATTGTTGAAGGAGATCT 241
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY  193 GGCATATAAGGGTTGCAACTATTATACGCTAATGGAACCTTTTATTGTTGAAGGAGATCT 252
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Db	28	ACAAATGGCTGTTTACAAAGTTAGTTCCTTGTCTCACCCTACTTCTTCTTGGAAATGTATCTA	87
QY	62	CTACTTGCAGAAACATGTTGATGCCAAGCGTGTGACTAGAGAAATGTGGTTCATTTTAGCTAT	121
Db	88	CTAGTAAGCACGCTGGACACAGCTAATGCTTGTACTAAAGAAATGTGGTAATCTTGGCTAT	147
QY	122	GGCATATGCCACGTTTCAGAAAGAGTCCCAAAAACCTATATGACCAATTTGTTGCTCA	191
Db	148	GGGATATGCCACGTTTCAGAAAGAGTCCAGAAAAATCCAAATATGTATGTAATTTGCTCT	207
QY	182	GGCTATAAGGGTTCACACTATTACAGTGTCTAAGGAGATTTGATTGTGAAGGAGATCT	241
Db	208	GGCTATAAGGGTTCACACTATTATAACGTAATGGAACCTTTTATTGTGAAGGAACGCTCT	267
QY	242	GACCTAGAAACCCAAAGATTTGACCTTCGAATGTGATACACAGATTTGCTTATTCAAAA	301
Db	268	GATCCAAAAATCCTAACATTTGCCCTCATATTTGTATGCCAATTTGCTATTCAAG	327
QY	302	TGTCCTGTTTCAGAAAGAAATGATAATTAACCCACCTGGATGACCACTTGTTCACG	361
Db	328	TGTCCACGTTTCAGAAAGAAAGAGATAATCTATCCACAGGATGTACGAGCTGTGCAC	387
QY	362	GGCTATACAGGTTGCTACTATTTCGATCAGATGGTATTTCCTCTGTAAGGAGAGAT	421
Db	388	GGTTACAAGGGTTGCTACTATTTTTGGTCAAGATGGAGAGTTTGTGTGAAGGAGAGT	447
QY	422	CCTGAACCCCAAG	433
Db	448	ATTGAACCTTAAG	459
RESULT 9			
BI433294			
LOCUS			
DEFINITION			
ESTS36055 P. infestans-challenged potato leaf, compatible reaction			
Solanum tuberosum cDNA clone PPCBC78 5' sequence, mRNA sequence.			
BI433294			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
1. 625			
/organism="Solanum tuberosum"			
/mol_type="mRNA"			
/cultivar="Kennebec"			
/db_xref="taxon:4113"			
/clone="PPBC78"			
/tissue_type="leaf"			
/dev_stage="6 week old"			
/lab_host="SOLR"			
/clone_lib="P. infestans-challenged potato leaf,			
compatible reaction"			
/notes="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:			
XhoI; supplier: Cornell University, Fry lab; sequencing:			
The Institute for Genomic Research; Whole plants were			

challenged with 20,000 sporangia/ml of the compatible P. infestans isolate US 940480. Leaf tissue was collected at 3, 6, 9, 12, 24, 48, 72 hours after inoculation and frozen in liquid nitrogen immediately upon removal. Kennebec plants showed first symptoms of infection at 48 hours after inoculation. NOTE: We cannot exclude the possibility that this sequence is actually derived from Phytophthora rather than potato."

Query Match	51.2%;	Score	270.8;	DB	4;	Length	625;
Best Local Similarity	79.7%;	Pred. No.	1.1e-53;				
Matches	349;	Conservative	0;	Mismatches	77;	Indels	12;
Gaps	2;						

QY	1	CATAATGGCTGTTTTCACA---AAGTTAGCTTTCCTTGCCTACTTGTCTTGGATGGAT	57
Db	23	CATAATGGCTGTTTTCACAAGGAAGTTAGTTTTCGTTTACCTACTAATTTGTTCTTGAAT	82
QY	58	GTTTCTACTTTGGCAACATGTTGATGCCAAGCGTTGTCTACTAGAGATGGTCAATTTTAG	117
Db	83	GTTTCTAT-----ATGTTGATGCTTTGGGTTGTCTAAAGAATGGTAAATCTTGG	133
QY	118	CTATGCGATATGCCACGTTTCAGAAGGAAGTCCCAAAACCTATATGACCAATTTGTTG	177
Db	134	CTTTGGGATATGCCACGTTTCAGAAGGAAGTCCGACAAATCCCATATGATCAATTTGTTG	193
QY	178	CTCAGGCTATAAGGGTTGCAACTATTACAGTGTCTAAAGGAGATTTGATTGTGAAGGAGA	237
Db	194	CTCAGGCTATAAGGGTTGTAATTTATATAGTGTCTTTCGGGAGATTTATTTGCGAAGGAGA	253
QY	238	ATCTGACCTTAGAACCCCAAGATTTGTAACCTTCGAATGTGTATACACAGATTCCTATTTC	297
Db	254	ATCTGACCCCAAAACCCCAAGCTTGCCTTAAATTTGTGATACAAATATTTCCTATTTC	313
QY	298	AAAAATGCTCTCTTCAGAGGAAGATGATAATTAACCCACCTGGATGACCACTTGTGTTG	357
Db	314	AGATGTCCTTTCAGAGGAAGAAATCGCTAATTTATCCACCGAATGTACCAATGCTG	373
QY	358	CACGGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGTATTTGTTGTGAAGGAGA	417
Db	374	CACAGGGTATAAGGATTTGCTACTATTTCGATAAAATGGAAGTTGTTGATGTGAAGGAGA	433
QY	418	GAGTCTGTAACCCCAAGAC	435
Db	434	GAGTGAATGAACCCCAAGGC	451

BI432721 628 bp mRNA linear EST 30-APR-2003
ESTS35482 P. infestans-challenged potato leaf, compatible reaction
Solanum tuberosum cDNA clone PPCAU32 5' sequence, mRNA sequence.

BI432721
EST.
Solanum tuberosum (potato)
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 628)
Restrepo,S., Griffiths,H.M., Smart,C.D., Cho,J., Chiemingo,A.,
Bougr,O., Buell,C.R., Ronning,C.M., Fry,W.E. and Baker,B.
Generation of ESTs from Potato Leaves Challenged with Phytophthora
infestans, Compatible Interaction
Unpublished (2000)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
<http://genome.arizona.edu/orders/>

RESULT 10
BI432721
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

Seq primer: M13P-R.
FEATURES
  source
    Location/Qualifiers
      1..628
        /organism="Solanum tuberosum"
        /mol_type="mRNA"
        /cultivar="Kennebec"
        /db_xref="taxon:4113"
        /clone="PPCAU32"
        /tissue_type="leaf"
        /dev_stage="6 week old"
        /lab_host="SOLR"
        /clone_lib="p. infestans-challenged potato leaf,
compatible reaction"
      /notes=vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Cornell University, Fry lab; sequencing:
The Institute for Genomic Research; whole plants were
challenged with 20,000 sporangia/ml of the compatible P.
infestans isolate US 940480. Leaf tissue was collected at
3, 6, 9, 12, 24, 48, 72 hours after inoculation and frozen
in liquid nitrogen immediately upon removal. Kennebec
plants showed first symptoms of infection at 48 hours
after inoculation. NOTE: We cannot exclude the
possibility that this sequence is actually derived from
Phytophthora rather than potato."
ORIGIN
  Query Match 51.2%; Score 270.8; DB 4; Length 628;
  Best Local Similarity 79.7%; Pred. No. 1.1e-63;
  Matches 349; Conservative 0; Mismatches 77; Indels 12; Gaps 2;

QY 1 CATATGGCGTGTTCACA---AAGTTAGCTTCCTTGGCTTGCCTACTTGTCTTGGATGGAT 57
DB 26 CATATGGCGTGTTCACAAGGAAGTATGTTTGGTGTCTTACTACTAATTTGTTCTTGGAA 85
QY 58 GTTCTTACTTGCAGAACATGTCATGCCAAGGCTTGTACTAGAGATGTTGTCATTTAG 117
DB 86 GTTCTTAT-----ATGTTGATGCTTTGGGTTGTACTAAAGAAATGTTAATCTTGG 136
QY 118 CTATGGCATATGCCACAGTTTCAGAAGGAAGTCCCAAAACCTTATATGACCAATTTGTTG 177
DB 137 CTTTGGGATATGCCACAGTTTCAGAAGGAAGTCCGACAAATCCCATATGATCAATTTGTTG 196
QY 178 CTCAGGCTATAGGGTTTGCACCTATTATACAGTGTCTAAAGGAGATTTGATTTGTGAAGGAG 237
DB 197 CTCAGGCTATAGGGTTTGTAAATTTATATAGTGTCTTCGGGAGATTTATTTTTCGAAGGAG 256
QY 238 ATCTGACCTTAGAACCCCAAGATTTGACCTTCGAATGTGTACACAGATTTGCTTATTC 297
DB 257 ATCTGACCCCAAAACCCCAAGCTTGCCTTAAATTTGTGATACAAATTTGCTTATTC 316
QY 298 AAAATGCTCTCGTTTCAGAAGGAAGATGATATTAATTAACCCACCTGGATGCACCACTTTGTTG 357
DB 317 AGATGTCCTCCGTTTCAGAGGAATTCGCTAATTTATCCACCGGATGTACCATGCTG 376
QY 358 CACGGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGATTTTGTCTGTGAAGGAG 417
DB 377 CACAGGATATAGGATTTGCTACTATTTCGATATAAAATGGCAAGTTTGTATGTGAAGGAG 436
QY 418 GAGTCCTGAACCAAGAC 435
DB 437 GAGTGATGAACCAAGGC 454

RESULT 11
BF153853 672 bp mRNA linear EST 11-SEP-2002
LOCUS 032C04 Mature tuber lambda ZAP Solanum tuberosum cDNA 5' similar to
DEFINITION PROTEINASE INHIBITOR TYPE II CM7 PRECURSOR . . . sp|Q43652, mRNA
sequence.
ACCESSION BF153853
VERSION BF153853.1 GI:11035793
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)

ORGANISM
  Solanum tuberosum
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  asterids; lamiids; Solanales; Solanaceae; Solanum.
  REFERENCE
  1 (bases 1 to 672)
  AUTHORS Crookshanks, M., Emmersen, J., Welinder, K.G. and Nielsen, K.L.
  TITLE The potato tuber transcriptome: analysis of 6077 expressed sequence
  tags
  JOURNAL FEBS Lett. 506 (2), 123-126 (2001)
  MEDLINE 21475600
  PUBMED 11591384
  COMMENT
    Contact: Karen G. Welinder
    Institut for bioteknologi
    Aalborg Universitet
    Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
    Tel: +45 96358467
    Fax: +45 98141808
    Email: kgw@bio.auc.dk
    High quality sequence stop: 672
    POLYA=yes.
  FEATURES
    Location/Qualifiers
      1..672
        /organism="Solanum tuberosum"
        /mol_type="mRNA"
        /cultivar="Field grown Kurag"
        /db_xref="taxon:4113"
        /tissue_type="tuber"
        /clone_lib="Mature tuber lambda ZAP"
        /note="Vector: Lambda ZAP"
  ORIGIN
    Query Match 50.9%; Score 269.4; DB 2; Length 672;
    Best Local Similarity 78.3%; Pred. No. 2.8e-63;
    Matches 353; Conservative 0; Mismatches 86; Indels 12; Gaps 2;

QY 1 CATATGGCGTGTTCACA---AAGTTAGCTTCCTTGGCTTGCCTACTTGTCTTGGATGGAT 57
DB 28 CATCATGGCTGTTCACAAGGAAGTAAATTTGTTGCTTACCTACTAATTTGTTCTTGGAA 87
QY 58 GTTCTTACTTGT-----CGAAACATGTTGATGCAAGGCTTGTACTAGAGAATTTG 108
DB 88 ATTCTACTTGTAGCGTGTGGACATGTTGATGCGAAGGCTTGTACTTTAGAATGTG 147
QY 109 TCATTTTACTATGGCATATGCCACAGTTTCAGAAGGAAGTCCCAAAACCTTATATGCAC 168
DB 148 TAACTCTTGGATTTGGGATATGCCACGTTTAGAAGGAAGTCCGACAAATCCCATATGCAT 207
QY 169 CAATTGTTCTCAGGCTATAGGGTTGCAACTATTACAGTGTCTAAAGGAGATTTGATTTG 228
DB 208 CAATTGTTCTCAGGCTATAGGGTTGTAATTTATATAGTGTCTTTTGGGAGATTTATTG 267
QY 229 TGAAGGAGAAATCTGACCCCTAGAAACCCAAAGAAATTTGACCTTCGAATGTGATACACAGAT 288
DB 268 TGAAGGAGAAATCTGACCCCAAAACCCAAAGAAATTTGACCTTCGAATGTGATACAAATAT 327
QY 289 TGCCTATTCAAAATGTCCTCGTTTCAGAAGGAAGATGATAATTAACCCACTGGATGCAC 348
DB 328 TGCCTATTCAAGATGTTCCCGTTTCAAAAGGAACAACTAATTTATCCACCGGATGTAC 387
QY 349 CACTTGTTCACGGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGTGATTTCTCTG 408
DB 388 CACATGCTGCACAGGGTACAGGGTTGCTACTATTTCGTTAAATTTGCGTAAATTTGCAAGTTGTATG 447
QY 409 TGAAGGAGAGAGTCTCTGAACCCCAAGACCACT 439
DB 448 TGAAGGAGAGAGTGTATGAGCCCCCAAGCAAAAT 478

RESULT 12
BI436259 675 bp mRNA linear EST 10-MAR-2003
LOCUS EST539020 cSTE Solanum tuberosum cDNA clone cSTE22115 5' sequence,
DEFINITION mRNA sequence.
```

```

ACCESSION BI436259
VERSION BI436259.1 GI:15260949
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 675)
AUTHORS van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R., Cho,J.,
Chiemingo,A., Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and
Baker,B.
TITLE Generation of ESTs from in vitro grown microtubers
JOURNAL Unpublished (2001)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13P-R.
FEATURES             Location/Qualifiers
     source            1..675
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Bintje"
                     /db_xref="taxon:4113"
                     /clone="cSTRE22115"
                     /tissue_type="axillary buds of stem explants; growing
                     sink-tubers"
                     /dev_stage="7, 8 and 10 days"
                     /lab_host="SOLR"
                     /clone_lib="cSTg"
                     /notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Tissue supplied by Christian Bachem and Richard
                     Visser (Department of Plant Breeding, Wageningen
                     University, The Netherlands). The cSTA libraries will
                     attempt to capture the induction and initiation/initial
                     growth of the tuber in an in vitro system as described in
                     Bachem et al. (Plant Journal, 1996). Small microtubers
                     develop from axillary buds attached to stem explants when
                     placed on a high sucrose medium (10%). Visible
                     morphological changes occur synchronously at day five in
                     the axillary buds. The first library, cSTA (1-20) consists
                     of axillary buds harvested on days 1-3. This targets
                     those genes involved in induction of the microtubers. The
                     following libraries, cSTA (21-40) and cSTA (41-60),
                     capture genes involved in tuber initiation and outgrowth.
                     This library is noted as P3 in Tanksley lab notebooks."

ORIGIN
Query Match      50.9%; Score 269.4; DB 4; Length 675;
Best Local Similarity 78.9%; Pred. No. 2.8e-63;
Matches 337; Conservative 0; Mismatches 81; Indels 9; Gaps 1;

QY 18 AGGTAGCTTCCTGCTTGGCTTACTTCTTCTCGATGGATGTTCTACTTT-----G 68
DB 28 AAGTTAGTTTCCTTGCTTACTTACTTCTTCTTGGATTATGCTACTTGTGAAGCAGG 87

QY 69 CGAAACATGTTGATGCCAAGCTTGCTACTAGAGAAATGGTCAATTTTAGCTATGGCATAT 128
DB 88 TGGACATGTTGATGCCAAGCTTGCTACTAAGAAATGGTGAATCTTGGCTATGGGATAT 147

QY 129 GCCACGTTCAGAGGAAGTCCCAAAACCTATATGACCAATTTGTTGCTCAGGCTATA 188
DB 148 GCCACGTTTCAAGAGGAAGTACGGAAATCCCATATGACCAACTGTTGTGCAAGGTATA 207

QY 189 AGGGTTGCACTATTACAGTGCCTAAAGGAGATTTGATTTGTGAAGGAGATCTGACCTTA 248
DB 208 AAGGTTGCAATTTATATAGTGCCTAAAGGAGATTTTCAATTTGCGAAGGACAACTGACCCAA 267

QY 249 GAAACCCCAAGAGATTGTACCTTCGAATGTGATACACAGATTTGCTTATTCAAAATGTCCTC 308

```

```

Db 268 AAAACCCAAAGCTTGCCTCCGAAATTTGATGCACACATTTGCTATTCAAAGTGTCTTC 327
QY 309 GTTTCAGAGGAAGATGATAATTAAACCCACCTGGATGCACCACTTGTTCACCGGGTATC 368
Db 328 GTTCCAGAGGAAGAGCGCTAATATATCCACGGGATGTACCACATGTTGCACAGGGTACA 387
QY 369 AGGGTTGCTACTATTATTTTCGATCAAGATGGTGATTTTCTGTGAAGGAGAGATCCTGAAC 428
Db 388 AGGATTGCTACTATTATTTTGGTAAAGACGGCAAGTTTGTGTGAAGGAGAGATTTGAAC 447
QY 429 CCNAGAC 435
Db 448 CCAAGGC 454

RESULT 13
BF153879
LOCUS
DEFINITION
ACCESSION BF153879
VERSION BF153879.1 GI:11035819
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 677)
AUTHORS Crookshanks,M., Emmersen,J., Welinder,K.G. and Nielsen,K.L.
TITLE The potato tuber transcriptome: analysis of 6077 expressed sequence
tags
JOURNAL FEBS Lett. 506 (2), 123-126 (2001)
MEDLINE 21475600
PUBMED 11591384
COMMENT Contact: Karen G. Welinder
Institut for bioteknologi
Aalborg Universitet
Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
Tel: +45 96356467
Fax: +45 98141808
Email: kgw@bio.auc.dk
High quality sequence stop: 677
POLYA=yes.
FEATURES             Location/Qualifiers
     source            1..677
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Field grown Kuras"
                     /db_xref="taxon:4113"
                     /tissue_type="Tuber"
                     /clone_lib="Mature tuber lambda ZAP"
                     /notes="Vector: Lambda ZAP"

ORIGIN
Query Match      50.9%; Score 269.4; DB 2; Length 677;
Best Local Similarity 78.3%; Pred. No. 2.8e-63;
Matches 353; Conservative 0; Mismatches 86; Indels 12; Gaps 2;

QY 1 CATATGGCGTGTTCACA---AAGTTAGCTTCCCTTGGCTTGCCTACTTGTCTTGGATGGAT 57
DB 28 CATCATGGCGTGTTCACAAGGAAGTTAATTCGTTGCTTACTACTAATTTGTTCTTGAAT 87

QY 58 GTTCTTACTTG-----CGAAACATGTTGATGCCAAGGCTTGTGTAGAGAAATGTGG 108
DB 88 ATTCTTACTTGTGTAGCGTGTGGAAATGTTGTGATGCGAAGGCTTGTACTTTTAGAATGTGG 147

QY 109 TCATTTTACTATGGCATATGCCACGTTTCAGAAGGAAGTCCCAAAACCTATATATGCAC 168
DB 148 TAATCTTGGATTGGGATATGCCCGCTTTAGAAGGAAGTCCGACAAATCCCATATGCAT 207

```

```

QY 169 CAATTGTTGCTCAGGCTATAGGGTTGCAACTATTACAGTGTAAAGGAGATTGATTTG 228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 208 CAATTGTTGCTCAGGCTATAGGGTTGTAATATTATAGTGTCTTTGGGAGATTATTTG 267
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 229 TGAAGAGAAATCTGACCTAGAACCCCAAGAGTTGTACCTTCGAATGTGATACACAGAT 288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 268 TGAAGAGAAATCTGACCCAAAACCCAAAGCTTGCCCTTAAATTGTGATACAAATAT 327
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 289 TGCTTATCAAAATGTCCTCGTTCAGAGGAAGAGATTAATTAACCCCACTGGATGCAC 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 328 TGCTTATCAAGATGTTCCCGTTCAAAAGGAAAAACACTAATTTATCCCAACCGGATGAC 387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 349 CACTTGTTCACGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGTATTGTCG 408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 388 CACATGCTGCACAGGGTACAGGGTTGCTACTATTTTCGGTAAATAATGGCAAGTTTGTATG 447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 409 TGAAGAGAGAGTCTGAAACCCCAAGACCACT 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 448 TGAAGAGAGAGTGATGAGCCCAAGGCAAT 478
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||

RESULT 14
BF153508          591 bp      mRNA      linear      EST 11-SEP-2002
LOCUS             027E11 Mature tuber lambda ZAP Solanum tuberosum cDNA 5' similar to
DEFINITION        PROTEINASE INHIBITOR TYPE II CM7 PRECURSOR . . . sp|Q43652. mRNA
sequence.
ACCESSION         BF153508
VERSION           BF153508.1 GI:11035448
KEYWORDS          Solanum tuberosum (potato)
SOURCE            Solanum tuberosum
ORGANISM          Solanum tuberosum

REFERENCE
AUTHORS           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                  asterids; lamids; Solanales; Solanaceae; Solanum.
TITLE             1 (bases 1 to 591)
AUTHORS           Crookshanks,M., Emmersen,J., Welinder,K.G. and Nielsen,K.L.
                  The potato tuber transcriptome: analysis of 6077 expressed sequence
                  tags
JOURNAL           FEBS Lett. 506 (2), 123-126 (2001)
MEDLINE           21475600
PUBMED            11591384
COMMENT           Contact: Karen G. Welinder
                  Institut for bioteknologi
                  Aalborg Universitet
                  Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
                  Tel: +45 96358467
                  Fax: +45 98141808
                  Email: Kgw@bio.auc.dk
                  High quality sequence stop: 591
                  POLYA=No.

FEATURES
source            1. .591
                  /organism="Solanum tuberosum"
                  /mol_type="mRNA"
                  /cultivar="Field grown Kuras"
                  /db_xref="taxon:4113"
                  /tissue_type="Tuber"
                  /clone_lib="Mature tuber lambda ZAP"
                  /note="Vector: Lambda ZAP"

ORIGIN
Query Match      50.4%; Score 267.8; DB 2; Length 591;
Best Local Similarity 78.0%; Pred. No. 7.5e-63;
Matches 352; Conservative 0; Mismatches 87; Indels 12; Gaps 2;

QY 1 CATAATGGCTGTTACACA---AAGTAGCTTCCTGCTGCTACTTGTCTTCCTGGATGGAT 57
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 9 CATCATGGCTGTTACAAAGGAAGTAAATTCGTTGCTTACCTACTAATTCGTTCTTCTGAAAT 68
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||

QY 58 GTTCTCTACTTG-----CGAAACATGTTGATGCCAAGGCTGTACTAGAAATGTGG 108
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 69 ATTTCTACTTGTGTAGCGGTGGAAACATGTTGATGCGAAGGCTTGTACTTGTAGAAATGTGG 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||

```

```

QY 109 TCATTTTAGCTATGCCATATGCCACGTTTCAGAGGAAGTCCCAAAAACCTATATATGCAC 168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 129 TAATCTTGGATTTGGGATATGCCACGTTCTAGAGGAAGTCCCAAAATCCCATATGCAT 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 169 CAATTGTTGCTCAGGCTATAAGGGTTTGCACACTATTACAGTGTCAAAGGAGATTGATTTG 228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 189 CAATTGTTGCTCAGGCTATAAGGGTTGTAATTAATTATAGTGTCTTTGGGAGATTATTTG 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 229 TGAAGAGAAATCTGACCTAGAACCCCAAGAGTTGTACCTTCGAATGTGATACACAGAT 288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 249 TGAAGAGAAATCTGACCCAAAACCCAAAGCTTGCCCTTAAATTGTGATACAAATAT 308
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 289 TGCTTATCAAAATGTCCTCGTTCAGAGGAAGAGATTAATTAACCCCACTGGATGCAC 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 309 TGCTTATCAAGATGTTCCCGTTCAAAAGGAAAAACACTAATTTATTCCTCCCAACCGGATGAC 368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 349 CACTTGTTCACGGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGTATTGTCGTG 408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 369 CACATGCTGCACAGGGTACAGGGTTGCTACTATTTCGGTAAATAATGGCAAGTTTGTATG 428
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
QY 409 TGAAGAGAGAGTCTTGAACCCCAAGACCACT 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||
Db 429 TGAAGAGAGAGTGATGAGCCCAAGGCAAT 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||

RESULT 15
BF1421162          779 bp      mRNA      linear      EST 16-AUG-2001
LOCUS             EST531828 tomato callus, TAMU Lycopersicon esculentum cDNA clone
DEFINITION        cLEC6619 5' end, mRNA sequence.
ACCESSION         BF1421162
VERSION           BF1421162.1 GI:15194442
KEYWORDS          Lycopersicon esculentum (tomato)
SOURCE            Lycopersicon esculentum
ORGANISM          Lycopersicon esculentum

REFERENCE
AUTHORS           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                  asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
TITLE             1 (bases 1 to 779)
AUTHORS           Alcalá,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
                  Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S.,
                  Romling,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
                  Giovannoni,J.
                  Generation of ESTs from tomato callus tissue
                  Unpublished (1999)
JOURNAL           Contact: CUGI
                  Clemson University Genomics Institute
                  Clemson University
                  100 Jordan Hall, Clemson, SC 29634, USA
                  Email: http://www.genome.clemson.edu/orders/index.html.

FEATURES
source            1. .779
                  /organism="Lycopersicon esculentum"
                  /mol_type="mRNA"
                  /cultivar="TA496"
                  /db_xref="taxon:4081"
                  /clone="cLEC6619"
                  /tissue_type="callus"
                  /dev_stage="25-40 days old"
                  /lab_host="XLI-Blue MRP"
                  /clone_lib="tomato callus, TAMU"
                  /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
                  XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
                  of seedlings 7-10 days post-germination were excised, cut
                  at both ends and placed on MS medium with no selection.
                  Mixed callus was harvested at 25 and 40 days and included
                  undifferentiated masses. Tomato Callus EST Library"

```

ORIGIN

Query Match 50.5%; Score 267.2; DB 4; Length 779;
 Best Local Similarity 82.9%; Pred. No. 1.2e-62;
 Matches 305; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy	68	CGCAAA	CAATGTTG	ATGCGCA	AGGCTTGT	ACTAGAG	AATGTG	GTCA	TTTTAG	CTATGG	CATA	127
Db	35	GTGGAA	CAATGTTG	ATGCGCA	AGGCTTGT	ACTAGAG	AATGTG	GTCA	TTTTAG	CTATGG	CATA	94
Qy	128	TGCCCA	CGCTTC	GAGAGG	AAAGTCC	CCAAAA	CCATAT	GCACCA	AAATGTTG	CTCAGG	CTAT	187
Db	95	TGCCCA	CGCTTC	GAGAGG	AAAGTCC	CCAAAA	CCATAT	GCACCA	AAATGTTG	CTCAGG	CTAT	154
Qy	188	AAGGTT	CGCAACTA	TATACAG	TGCTAA	AGGAGAT	TTGATT	TGTGA	AGGAGAA	TCTGAC	CCCT	247
Db	155	AAAGTT	CGCAATTA	TATAGT	GTAA	TGGGAC	TTTAT	TGTGA	AGGACAA	TCTGAC	CCCA	214
Qy	248	AGAAAC	CCAAAG	ATGTG	ACCTTC	GAAATG	TGTAT	CACAG	ATTGCTTA	TATCAAA	ATGTCT	307
Db	215	AGAAAC	CCAAAC	CTTGC	CCCTTAA	ATTGTG	ATCC	ACATAT	TTCGCT	ATTCAAA	GTGTCC	274
Qy	308	CGTTCAG	AGGAAAG	ATGAT	TAATTA	AAACCA	CTGGAT	GCACCA	CACTTGT	TGACGGG	CTAT	367
Db	275	CGTTCAG	AGGAAAG	ATGAT	TAATTA	AAACCA	CTGGAT	GCACCA	CACTTGT	TGACGGG	CTAT	334
Qy	368	CAGGTT	CGCTACTA	TTTCCG	ATCGT	TGATTT	TGCTGT	GTAAGG	AGAGAG	TCTCTG	AA	427
Db	335	AAAGTT	CGCTACTA	TTTCGG	TAAAG	ATGCGC	AGTTT	GTTTGT	GTAAGG	AGAGAT	GTATG	394
Qy	428	CCCAAG	AC	435								
Db	395	CCCAAG	GC	402								

Search completed: January 22, 2005, 07:16:48
Job time : 2589 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 05:24:58 ; Search time 439 Seconds

(without alignments)
6923.861 Million cell updates/sec

Title: US-10-725-829-1

Perfect score: 529

Sequence: 1 cataatggctgttcacaag.....ctgtaatggactttatttg 529

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0 .

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	529	100.0	529	18	US-10-725-829-1
2	288.4	54.5	692	18	US-10-725-829-3
3	178.8	33.8	1360	14	US-09-812-502-2
4	178.8	33.8	1360	14	US-10-164-961-2
5	178.8	33.8	1360	15	US-10-157-622-2
6	149	28.2	1104	14	US-09-812-502-1
7	149	28.2	1104	14	US-10-164-961-1
8	149	28.2	1104	15	US-10-157-622-1
9	149	28.2	1104	15	US-10-072-809A-56
10	35.6	6.7	680	13	US-10-027-632-281276
11	35.6	6.7	680	15	US-10-027-632-281276
12	35.6	6.7	108359	14	US-10-191-807-3

Sequence 1, Appli
Sequence 3, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 56, Appli
Sequence 281276,
Sequence 281276,
Sequence 3, Appli

c	13	35.2	6.7	6134	16	US-10-221-714A-285	Sequence 285, App
c	14	35.2	6.7	6850	15	US-10-311-455-611	Sequence 611, App
c	15	35.2	6.7	10433	15	US-10-311-455-352	Sequence 352, App
c	16	35	6.6	9964	15	US-10-311-455-71	Sequence 71, Appl
c	17	34.8	6.6	632	13	US-10-027-632-96703	Sequence 96703, A
c	18	34.8	6.6	632	15	US-10-027-632-96703	Sequence 96703, A
c	19	34.8	6.6	6866	15	US-10-204-708-19	Sequence 19, Appl
c	20	34.8	6.6	6866	15	US-10-311-455-639	Sequence 639, App
c	21	34.8	6.6	152330	13	US-10-087-192-1834	Sequence 1834, App
c	22	34.6	6.5	346	9	US-09-880-107-860	Sequence 860, App
c	23	34.4	6.5	789	15	US-10-369-493-42185	Sequence 42185, A
c	24	34.4	6.5	63155	15	US-10-292-798-449	Sequence 449, App
c	25	34.4	6.5	193303	15	US-10-081-327-37	Sequence 37, Appl
c	26	34.4	6.5	193303	15	US-10-081-327-44	Sequence 44, Appl
c	27	34.2	6.5	1262	9	US-09-070-927A-759	Sequence 759, App
c	28	34.2	6.5	1479	16	US-10-425-114-36514	Sequence 36514, A
c	29	34.2	6.5	1479	18	US-10-425-115-22922	Sequence 22922, A
c	30	34	6.4	32249	10	US-09-764-891-7477	Sequence 7477, Ap
c	31	34	6.4	75899	9	US-09-854-883-243	Sequence 243, App
c	32	34	6.4	75899	15	US-10-360-510-243	Sequence 243, App
c	33	34	6.4	518360	17	US-10-367-094-125	Sequence 125, App
c	34	33.8	6.4	1059	13	US-10-027-632-250275	Sequence 250275,
c	35	33.8	6.4	1059	15	US-10-027-632-250275	Sequence 250275,
c	36	33.8	6.4	5473	15	US-10-037-270-260	Sequence 260, App
c	37	33.8	6.4	5473	15	US-10-117-722-260	Sequence 260, App
c	38	33.8	6.4	7545	17	US-10-473-575-25	Sequence 25, Appl
c	39	33.4	6.3	255	10	US-09-930-213-697	Sequence 697, App
c	40	33.4	6.3	648	9	US-09-861-451A-25	Sequence 25, Appl
c	41	33.4	6.3	668	13	US-10-027-632-144414	Sequence 144414,
c	42	33.4	6.3	668	15	US-10-027-632-144414	Sequence 144414,
c	43	33.4	6.3	2447	16	US-10-264-237-715	Sequence 715, App
c	44	33.4	6.3	39560	15	US-10-085-117-118	Sequence 118, App
c	45	33.4	6.3	96597	11	US-09-997-722-289	Sequence 289, App

ALIGNMENTS

RESULT 1

US-10-725-829-1
; Sequence 1, Application US/10725829
; Publication No. US20040205846A1
; GENERAL INFORMATION:
; APPLICANT: Chye, Mee Len
; APPLICANT: Xu, Zeng-Fu
; APPLICANT: Sin, Suk Fong
; TITLE OF INVENTION: Genetically Modified Plants with Heterogenous Proteinase Inhibitor
; TITLE OF INVENTION: SapiN2A or SapiN2B and Methods of Use Thereof for the Inhibition
; TITLE OF INVENTION: Of Trypsin- and Chymotrypsin-Like Activities
; FILE REFERENCE: 9661-043-999
; CURRENT APPLICATION NUMBER: US/10/725,829
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: 60/429,992
; PRIOR FILING DATE: 2002-11-29
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Solanum americanum
US-10-725-829-1

Query Match 100.0%; Score 529; DB 18; Length 529;
Best Local Similarity 100.0%; Pred. No. 3.8e-142;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATAATGGCTGTTCACAAAGTTAGCTTCCTTGGCTGCTACTTGTCTTGGATGGATGTT 60

Db 1 CATAATGGCTGTTCACAAAGTTAGCTTCCTTGGCTGCTACTTGTCTTGGATGGATGTT 60

Qy 61 TCTACTTGGCAACATGTTGATGCCAAGGCTTCTACTAGAGAATGTTGTCATTTAGCTA 120

Db 61 TCTACTTGGCAACATGTTGATGCCAAGGCTTCTACTAGAGAATGTTGTCATTTAGCTA 120

```
QY 121 TGGCATATGCCACGTTTCAGAGGAAGTCCCAAAACCTATATGACCAATTTGTTGCTC 180
Db 121 TGGCATATGCCACGTTTCAGAGGAAGTCCCAAAACCTATATGACCAATTTGTTGCTC 180
QY 181 AGGCTATAAGGGTTGCAACTATTACAGTGTCTAAAGGAGATTGATTTGTGAAGAGAATC 240
Db 181 AGGCTATAAGGGTTGCAACTATTACAGTGTCTAAAGGAGATTGATTTGTGAAGAGAATC 240
QY 241 TGAACCTAGAACCCAAAGAGATTGTACCTTCGATGTGATACACAGATTGCTTATTCMAA 300
Db 241 TGAACCTAGAACCCAAAGAGATTGTACCTTCGATGTGATACACAGATTGCTTATTCMAA 300
QY 301 ATGCTCTCGTTTCAGAGGAAGATGATAATTAACACCACTGGATGCACCACTTTGTTGCAC 360
Db 301 ATGCTCTCGTTTCAGAGGAAGATGATAATTAACACCACTGGATGCACCACTTTGTTGCAC 360
QY 361 GGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGAATTTTGTCTGTGAAGGAGAG 420
Db 361 GGGCTATCAGGGTTGCTACTATTTCGATCAAGATGGTGAATTTTGTCTGTGAAGGAGAG 420
QY 421 TCCGTGAACCAAGACCACTGCTTATTCTTAATCAATCATATGTTGTTATCTATCAAAAAA 480
Db 421 TCCGTGAACCAAGACCACTGCTTATTCTTAATCAATCATATGTTGTTATCTATCAAAAAA 480
QY 481 AAATATGTTATGCATGATATATGCTGTTACTGTAAATGGACATTTATTG 529
Db 481 AAATATGTTATGCATGATATATGCTGTTACTGTAAATGGACATTTATTG 529

RESULT 2
US-10-725-829-3
; Sequence 3, Application US/10725829
; Publication No. US20040205846A1
; GENERAL INFORMATION:
; APPLICANT: Chye, Mee Len
; APPLICANT: Xu, Zeng-Fu
; APPLICANT: Sin, Suk Fong
; TITLE OF INVENTION: Genetically Modified Plants with Heterogenous Proteinase Inhibitor
; TITLE OF INVENTION: SAPIZA or SAPIZB and Methods of Use Thereof for the Inhibition
; FILE REFERENCE: 9661-043-999
; CURRENT APPLICATION NUMBER: US/10/725,829
; PRIOR FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: 60/429,992
; PRIOR FILING DATE: 2002-11-29
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 692
; TYPE: DNA
; ORGANISM: Solanum americanum
US-10-725-829-3

Query Match 54.5%; Score 288.4; DB 18; Length 692;
Best Local Similarity 80.5%; Pred. No. 1.3e-72;
Matches 352; Conservative 0; Mismatches 76; Indels 9; Gaps 1;

QY 18 AAGTTAGCTTCCTTGCTGCTACTGTTCTTCGATGGATGTTTCTACTTTGCG------ 70
Db 49 AAGTTAGTTCCCTTGCTTACTACTTGTCTTGGATTAAATGTTCTTACATGTAAGCGCG 108
QY 71 --AAACATGTTGATGTCNAGGCTTGATAGAGAAATGTCATTTAGCTATGGCATAT 128
Db 109 TAAAAACATGTTGATGTCNAGGCTTGATAGAGAAATGTCATTTAGCTATGGGAATAT 168
QY 129 GCCCAGCTTCAGAGGAAGTCCCAAAACCTATATGCACCAATTTGTTGCTCAGGCTATA 188
Db 169 GCCCGGTTTCAGAGGAAGTCCCGAAAATCCCATATATGCAGAAATTTGTTGCTCAGGCTATA 228
QY 189 AGGGTTGCAACTATTACAGTGTCTAAAGGAGATTGATTTGTGAAGGAGAAATCTGACCCCTA 248
Db 229 AAGGTTGCAACTATTATAGTGTCTAAATGGGACTTTTATTTTTCGAGGAAGTCTCTGACCCCTA 288
```

```
QY 249 GAAACCCAAAAGATTGTACCTTCGAATGTGATACACAGATTGCTTATTCAAAAATGTCCTC 308
Db 289 AAAACCCAAAATACTTCCCTCTTATTTTGTGTGAGATATTTGCTATTCAAAATGTCCTCC 348
QY 309 GTTCAGAGGAAGATGATAATTAACCCACCTGGATGCACCACTTGTTCACCGGGCTATC 368
Db 349 GTTCAGAGGAAGACTATAATATATATCCACCGGATGCACCACTTGTTCACCGGGGTACA 408
QY 369 AGGGTTGCTACTATTTCGATCAAGATGTGATTTTCTGTGTGAAGGAGAGAGTCCCTGAAC 428
Db 409 AGGGTTGCTACTATTTTAGTAAAGAGGTGAGTTTGTGTGTGAAGGAGAGATGATGAAC 468
QY 429 CCAAGACCACTGCTTAT 445
Db 469 CCAACGTTATTCTAAT 485

RESULT 3
US-09-812-502-2
; Sequence 2, Application US/09812502
; Publication No. US20030027303A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; TITLE OF INVENTION: SEQUENCES ENCODING SAME
; FILE REFERENCE: 9748B
; CURRENT APPLICATION NUMBER: US/09/812,502
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US/09/431,500
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1360
; TYPE: DNA
; ORGANISM: Nicotiana glauca
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (97)...(1200)
US-09-812-502-2

Query Match 33.8%; Score 178.8; DB 10; Length 1360;
Best Local Similarity 67.3%; Pred. No. 7.9e-41;
Matches 303; Conservative 0; Mismatches 132; Indels 15; Gaps 3;

QY 5 ATGGCTGTTTCACAAAGTTAGCTTCCTTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTA 64
Db 10 ATGGCTGTTTCACAGAGTTAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 69
QY 65 CT-----TGCAGAACATGTTGATGCCAAGGCTTGTACTAGAGAAATGTGTGTCATT---TT 115
Db 70 GTAAGCAATGTGGAACATGCGATGCGAAGGCTTGTACCTTAACTGTGATCCAGAAAT 129
QY 116 AGCTATGCGATATGCCACGTTTCAGAGGAAGTCCCAAAACCTATATGACCAATTTGT 175
Db 130 GCCTATGAGGTTTGCCTGCTTTCAGAGGAAGAAAGAAATGATCGGATATGACCAACTGT 189
QY 176 TGCTCAGGCTATAGGGTTGCAACTATTACAGTGTCTAAAGGAGATTTGATTTCTGGAAGGA 235
Db 190 TCGCAGGCAAGGAGGTTGTAGTACTTTCAGTGTGATGATGATGATGATGATGATGATGATGAT 249
QY 236 GAATCTGACCTTAGAAACCCAAAGATTTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 295
Db 250 GAGTCTGATCTTAGAAATCCAAAGGCTTGTACCTTAACTGTGATCCAGAAATTTGCTAT 309
QY 296 TCAAAATGCTCTGTTTCAGAGGAAGATGATAATTTAAACCCACTGGATGACCACTTGT 355
```

Db 310 GGAGTTTCCCGCTTCAGAGAAAGAG-----AATGATCGGATATGCACCAACTGT 363
Qy 356 TGCACGGGTATCAGGGTCTACTATTTCGATCAAGATGGTGAATTTTGTCTGTGAAGGA 415
Db 364 TGGCAGGCACGAGAGGGTTGTAAGTACTTCAGTGATGATGGAACCTTTTGTGTGAAGGA 423
Qy 416 GAGAGTCTGACCCCAAGCAACTGCTTAT 445
Db 424 GAGTCTGATCTAGAAATCCAAAGGCTTGT 453

RESULT 4
US-10-164-961-2
; Sequence 2, Application US/10164961
; Publication No. US20030096388A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; Atkinson, Angela H.
; Heath, Robyn L.
; Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/164,961
; FILING DATE: 07-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-Sep-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 97..1200
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-164-961-2

Query Match 33.8%; Score 178.8; DB 14; Length 1360;
Best Local Similarity 67.3%; Pred. No. 7.9e-41;
Matches 303; Conservative 0; Mismatches 132; Indels 15; Gaps 3;

Qy 5 ATGGCTGTTCAAAAGTACTGCTTCTGCTGCTACTTGTTCGATGGAATGTTTCTA 64
Db 10 ATGGCTGTTCAAGAGTACTGCTTCTGCTGCTACTTGTTCGATGGAATGTTTCTGCT 69
Qy 65 CT-----TGCAGAAATGTTGATGCAAGGCTTGTACTAGAAATGTTGTCATT---TT 115
Db 70 GTAAGCAATGTGGAACATGCAGATGCCAAGGCTTGTACCTTAACTTGTGATCCAGAAAT 129

Qy 116 AGCTATGGCATATGCCACAGCTTCAGAGGAAGTCCCCAAAACCTATATGACCAACTTGT 175
Db 130 GCCTATGGAGTTTGCCTCGGCTTCAGAGAGAAAGAAATGATCGGATATGACCAACTGT 189
Qy 176 TGCTCAGGCTATAAGGGTTGCAACTATTACAGTGTCTAAAGGAGATTTGATTTGTGAAGGA 235
Db 190 TGGCAGGCACGAGAGGGTTGTAAGTACTTCAGTGATGATGGAACCTTTTGTGTGAAGGA 249
Qy 236 GAATCTGACCTAGAAACCCAAAGATTTGTACCTTCGAATGTGATACACAGATTTGCTTAT 295
Db 250 GAGTCTGATCTAGAAATCCAAAGGCTTGTACCTTAAACTGTGATCCAAAGAAATTCCTAT 309
Qy 296 TCAAAATGCTCGTTTCAGAGGAAGATGATTAATTAACCCACTGATGCACCACTTGT 355
Db 310 GGAGTTTGCCTCGCTTCAGAGGAAGATGATTAATTAACCCACTGATGCACCAACTGT 363
Qy 356 TGCACGGGTATCAGGGTTCCTACTATTTCGATCAAGATGGTGAATTTTGTCTGTGAAGGA 415
Db 364 TGGCAGGCACGAGAGGGTTGTAAGTACTTCAGTGATGATGGAACCTTTTGTGTGAAGGA 423
Qy 416 GAGAGTCTGAAACCCCAAGCAACTGCTTAT 445
Db 424 GAGTCTGATCTAGAAATCCAAAGGCTTGT 453

RESULT 5
US-10-157-622-2
; Sequence 2, Application US/10157622
; Publication No. US20030129720A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; Atkinson, Angela H.
; Heath, Robyn L.
; Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/157,622
; FILING DATE: 29-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-Sep-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS

Db	1	AAGGCTTGTACCTTAAACTGTGTATCCCAAGAAATTCCTTATGGAGTTTTCGCCCGCGTTTCAGAA	60		
QY	143	GGAGTCCCAAAAACCTATATATGACCAATTTGTTGCTCAGCTATATAGGGTTTGCACACTAT	202		
Db	61	GAAGAAGAATGATCGGATATGACCAACTGTTTGGCAGGACGAAAGGGTTGTAAAGTAC	120		
QY	203	TACAGTGTCTAAAGGAGATTTGATTTTGTGAAGGAGAATCTGACCTCTAGAAACCCCAAAAGAT	262		
Db	121	TTCACTGATGATGGAACCTTTTGTGTTGTAAGGAGAGTCTGATCCTAGAAATCCAAAGGCT	180		
QY	263	TGTACTCTCGAATGTGATACACAGATTCGCTTATTTCAAAAATGTCCTCGTTTCAGAAAGGAAAG	322		
Db	181	TGTACTTTAACTGTGATCCAAAGATTCGCTATGAGTGTTCGCCCGCTTCAGAAAGAAAG	240		
QY	323	ATGATAATTAACCCACTGGATGACCACTTTGTTGCA CGGCTATACGGGTTGCTACTAT	382		
Db	241	AAG-----AATGATCGGATATGACCAACTGTTTGGCAGGACGAAAGGGTTGTAAAGTAC	294		
QY	393	TTGATCAAGATGGTATTTTGTCTGTGAAGGAGAGTCTTGAACCCCAAGACCACTGCT	442		
Db	295	TTCACTGATGATGGAACCTTTTGTGTTGTAAGGAGAGTCTGATCCTAGAAATCCAAAGGCT	354		
QY	443	TAT 445			
Db	355	TGT 357			
RESULT 7					
US-10-164-961-1					
; Sequence 1, Application US/10164961					
; Publication No. US20030096388A1					
; GENERAL INFORMATION:					
; APPLICANT: Anderson, Marilyn A.					
; Atkinson, Angela H.					
; Heath, Robyn L.					
; Clarke, Adrienne E.					
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC					
; NUMBER OF SEQUENCES: 14					
; CORRESPONDENCE ADDRESS:					
; ADDRESSEE: Scully, Scott, Murphy & Presser					
; STREET: 400 Garden City Plaza					
; CITY: Garden City					
; STATE: New York					
; COUNTRY: United States of America					
; ZIP: 11530					
; COMPUTER READABLE FORM:					
; MEDIUM TYPE: Floppy disk					
; COMPUTER: IBM PC compatible					
; OPERATING SYSTEM: PC-DOS/MS-DOS					
; SOFTWARE: Patent In Release #1.0, Version #1.30					
; CURRENT APPLICATION DATA:					
; APPLICATION NUMBER: US/10/164,961					
; FILING DATE: 07-Jun-2002					
; CLASSIFICATION: <Unknown>					
; PRIOR APPLICATION DATA:					
; APPLICATION NUMBER: US/08/454,295					
; FILING DATE: 01-SEP-1995					
; ATTORNEY/AGENT INFORMATION:					
; NAME: Digiglio, Frank S.					
; REGISTRATION NUMBER: 31,346					
; REFERENCE/DOCKET NUMBER: 9748					
; TELECOMMUNICATION INFORMATION:					
; TELEPHONE: (516) 742-4343					
; TELEFAX: (516) 742-4366					
; TELEX: 230 901 SANS UR					
; INFORMATION FOR SEQ ID NO: 1:					
; SEQUENCE CHARACTERISTICS:					
; LENGTH: 1104 base pairs					
; TYPE: nucleic acid					
; STRANDEDNESS: single					
; TOPOLOGY: linear					
; MOLECULE TYPE: DNA (genomic)					
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:					

US-10-164-961-1

Query Match 28.2%; Score 149; DB 14; Length 1104;
Best Local Similarity 67.2%; Pred. No. 2.9e-32;
Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

QY 86 AAGCGTTGTACTAGAGAACTGTGTC---ATTTAGCTATGGCATATGCCACCGTTCAGAA 142
DB 1 AAGCGTTGTACTAGAGAACTGTGTC---ATTTAGCTATGGCATATGCCACCGTTCAGAA 60

QY 143 GGAAGTCCCCAAAACCTATATGACCAATTTGTTGCTCAGGCTATAGGGTTGCAACTAT 202
DB 61 GAAAGAGAGATGATCGGATATGACCAACTGTTGCCAGGCACGAGGGTTGTAAGTAC 120

QY 203 TACAGTGTCTAAGAGAGATTTGTTGTAAGAGAGATCTGACCTCTAGAAACCCAAAGAT 262
DB 121 TACAGTGTCTAAGAGAGATTTGTTGTAAGAGAGATCTGACCTCTAGAAACCCAAAGAT 180

QY 263 TGTACCTTCGAATGTATACACAGATGCTTATCAAAATGTTCTGTTGTAAGAGAGAT 322
DB 181 TGTACCTTCGAATGTATACACAGATGCTTATCAAAATGTTCTGTTGTAAGAGAGAT 240

QY 323 ATGATAAATAAACCCACTGATGACCACTGTTTGCACGGGCTATCAGGGTTGCTACTAT 382
DB 241 AAG-----ATGATCGGATATGACCAACTGTTGCCAGGCACGAGGGTTGTAAGTAC 294

QY 383 TTCGATCAAGATGGTATTTGTTGTAAGAGAGAGATCTGAAACCCAGACCACTGCT 442
DB 295 TTCGATCAAGATGGTATTTGTTGTAAGAGAGAGATCTGAAACCCAGACCACTGCT 354

QY 443 TAT 445
DB 355 TGT 357

RESULT 8

US-10-157-622-1
; Sequence 1, Application US/10157622
; Publication No. US20030129720A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; Atkinson, Angela H.
; Heath, Robyn L.
; Clarke, Adrienne E.

; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/157,622
; FILING DATE: 29-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1104 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-157-622-1

Query Match 28.2%; Score 149; DB 15; Length 1104;
Best Local Similarity 67.2%; Pred. No. 2.9e-32;
Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

QY 86 AAGCGTTGTACTAGAGAACTGTGTC---ATTTAGCTATGGCATATGCCACCGTTCAGAA 142
DB 1 AAGCGTTGTACTAGAGAACTGTGTC---ATTTAGCTATGGCATATGCCACCGTTCAGAA 60

QY 143 GGAAGTCCCCAAAACCTATATGACCAATTTGTTGCTCAGGCTATAGGGTTGCAACTAT 202
DB 61 GAAAGAGAGATGATCGGATATGACCAACTGTTGCCAGGCACGAGGGTTGTAAGTAC 120

QY 203 TACAGTGTCTAAGAGAGATTTGTTGTAAGAGAGATCTGACCTCTAGAAACCCAAAGAT 262
DB 121 TACAGTGTCTAAGAGAGATTTGTTGTAAGAGAGATCTGACCTCTAGAAACCCAAAGAT 180

QY 263 TGTACCTTCGAATGTATACACAGATGCTTATCAAAATGTTCTGTTGTAAGAGAGAT 322
DB 181 TGTACCTTCGAATGTATACACAGATGCTTATCAAAATGTTCTGTTGTAAGAGAGAT 240

QY 323 ATGATAAATAAACCCACTGATGACCACTGTTTGCACGGGCTATCAGGGTTGCTACTAT 382
DB 241 AAG-----ATGATCGGATATGACCAACTGTTGCCAGGCACGAGGGTTGTAAGTAC 294

QY 383 TTCGATCAAGATGGTATTTGTTGTAAGAGAGAGATCTGAAACCCAGACCACTGCT 442
DB 295 TTCGATCAAGATGGTATTTGTTGTAAGAGAGAGATCTGAAACCCAGACCACTGCT 354

QY 443 TAT 445
DB 355 TGT 357

RESULT 9

US-10-072-809A-56
; Sequence 56, Application US/10072809A
; Publication No. US20030217382A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn, A., Lay, Fung T., Heath, Robyn, L.
; TITLE OF INVENTION: Plant-derived molecules and genetic sequences encoding same and us
; TITLE OF INVENTION: therefor
; FILE REFERENCE: 18-01
; CURRENT APPLICATION NUMBER: US/10/072,809A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: USSN 60/267,271
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 56
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Nicotiana glauca
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1104)
US-10-072-809A-56

Query Match 28.2%; Score 149; DB 15; Length 1104;
Best Local Similarity 67.2%; Pred. No. 2.9e-32;
Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

QY 86 AAGCGTTGTACTAGAGAACTGTGTC---ATTTAGCTATGGCATATGCCACCGTTCAGAA 142
DB 1 AAGCGTTGTACTAGAGAACTGTGTC---ATTTAGCTATGGCATATGCCACCGTTCAGAA 60

```
Db      1  AAGGCTTGTTACTTAACCTGTGATCCAAAGAAATTGCCCTATGGAGTTTGGCCCGCGCTTCAGAA 60
QY      143  GGAAGTCCCCAAAACCTATATGACCAAAATGTTGCTCAGGCTATAAGGGTTGCAACTAT 202
Db      61  GAAAGAGAGATGATCGGATATGACCACTGTTGGCGGACGACGAGGGTTGTAAGTAC 120
QY      203  TACAGTGTAAAGGAGATTGATTGTGAAGGAGAATCTGACCCCTAGAAACCCAAAAGAT 262
Db      121  TTCAGTGTATGGAACCTTTTGTGTTGAAGGAGAGTCTGATCCTAGAAATCCAAAGGCT 180
QY      263  TGTACCTTCGAATGTATACACAGATTGCTTATCAAAATGTCTCGTTTCAGAGGAAG 322
Db      181  TGTACCTTAAACTGTGATCAAGAAATGCCATATGGAAGTTTGGCCCGGCTTCAGAGGAAG 240
QY      323  ATGATAAATAAACCCACTGATGCACCACTGTTTGCACGGGCTATCAGGGTTGCTACTAT 382
Db      241  AAG-----AATGATCGGATATGACCACTGTTGGCGGACGACGAGGGTTGTAAGTAC 294
QY      383  TTCGATCAAGATGGTATTTGTTGTGAAGGAGAGTCTCTGAACCCCAAGACCACTGCT 442
Db      295  TTCAGTGTATGGAACCTTTTGTGTTGAAGGAGAGTCTGATCCTAGAAATCCAAAGGCT 354
QY      443  TAT 445
Db      355  TGT 357

RESULT 10
US-10-027-632-281276
; Sequence 281276, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281276
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-281276
Query Match      6.7%; Score 35.6; DB 13; Length 680;
Best Local Similarity 53.6%; Pred. No. 12;
Matches 74; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY      246  CTGAAACCCAAAGATTGACCTTGAATGTGATACACAGATTGCTTATTCAAAATGTC 305
Db      27  CTTGACACCTAAGAGATTGCTTCTTGGAAATAAATACAGGTTTAGGTATTTAGATGAC 86
QY      306  CTCGTTTCAGAGGAAGATGATTAATTAACCCACTGATGCACCACTTGTTCACGGGCT 365
Db      87  CTAATTAAGTATATACCAAGTTATGAGAGGCTTGGTAAGCGGCTTGTGATGATGCTT 146
QY      366  ATCAGGGTTGCTACTATT 383
Db      147  TTAAGTTTACACCTTTT 164

RESULT 12
US-10-191-807-3/c
; Sequence 3, Application US/10191807
; Publication No. US20030068691A1
; GENERAL INFORMATION:
; APPLICANT: HU, Song et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL001275-PROV
; CURRENT APPLICATION NUMBER: US/10/191,807
; CURRENT FILING DATE: 2002-07-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 108359
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
```

```
Db      147  TTAAGTTTACACCTTTT 164

RESULT 11
US-10-027-632-281276
; Sequence 281276, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281276
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-281276
Query Match      6.7%; Score 35.6; DB 15; Length 680;
Best Local Similarity 53.6%; Pred. No. 12;
Matches 74; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY      246  CTGAAACCCAAAGATTGACCTTGAATGTGATACACAGATTGCTTATTCAAAATGTC 305
Db      27  CTTGACACCTAAGAGATTGCTTCTTGGAAATAAATACAGGTTTAGGTATTTAGATGAC 86
QY      306  CTCGTTTCAGAGGAAGATGATTAATTAACCCACTGATGCACCACTTGTTCACGGGCT 365
Db      87  CTAATTAAGTATATACCAAGTTATGAGAGGCTTGGTAAGCGGCTTGTGATGATGCTT 146
QY      366  ATCAGGGTTGCTACTATT 383
Db      147  TTAAGTTTACACCTTTT 164

RESULT 12
US-10-191-807-3/c
; Sequence 3, Application US/10191807
; Publication No. US20030068691A1
; GENERAL INFORMATION:
; APPLICANT: HU, Song et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL001275-PROV
; CURRENT APPLICATION NUMBER: US/10/191,807
; CURRENT FILING DATE: 2002-07-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 108359
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (1)...(108359)
; OTHER INFORMATION: n = A,T,C or G
US-10-191-807-3

Query Match      6.7%; Score 35.6; DB 14; Length 108359;
Best Local Similarity 45.9%; Pred. No. 1.6e+02;
Matches 122; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 235 AGAATCTGACCCCTAGAACCCAAAGATTGCTACCTTCGAAATGTGATACACAGATGCTTA 294
Db 850 ACAATATAACATAAATATATATTAATGATATATACCAACATATACACACATAGCATG 791
QY 295 TTCAAAATGTCCTCGTTTCAGAACGAAGATGATAATTAACCCACTGGATGCACCACTTG 354
Db 790 TGAAGGTATAAAAGATTACCTGAGTATGATAAATACCAATCAGGATAATGGTTAAAGG 731
QY 355 TTCACGGGCTATCAGGGTGTCTACTATTTTCGATCAAGATGGTGTGTTGCTGTCGAAG 414
Db 730 TTTTCAGTGTGTTAAATGGTATTTAAATTTCTTAATCTAGGTGATATTTATTTCTTAATCT 671
QY 415 AGAGAGTCTGTAACCCCAAGACCACTGCTTATTTCTAATCAATCATATGTTGTTATCTATC 474
Db 670 AATGAACATATATACATGAATGTTTCATATATATTTCTTTTATACTTCTTATATGTT 611
QY 475 AAAAAAATATGATGTCATGATATA 500
Db 610 CAAATTTTCTCTAATACTTTTTTTA 585

RESULT 13
US-10-221-714A-285/c
; Sequence 285, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: tumor suppressor genes and oncogenes
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221,714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 285
; LENGTH: 6134
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-285

Query Match      6.7%; Score 35.2; DB 16; Length 6134;
Best Local Similarity 44.9%; Pred. No. 49;
Matches 133; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 220 TTTGATTTGTGAAGGAGAACTCAGCCCTAGAAACCCCAAGATTGTACCTTCGAATGTGA 279
Db 4083 TTCGAAATTTTACAAATAAATAAATGAATTTAAACTTAAATATATAAATTTTAAACCTTA 4024
QY 280 TACACAGATTGCTTTATTCAAAATGTCTCTCGTTCAGAGGAAGATGATATTAACCCAC 339
```

```
Db 4023 ATATAAATTTACATCTACATATATACTCAATCAAAAAAACAACAAATATAAT 3964
QY 340 TGGATGCACCACCTTGTTCACGGGCTATCAGGGTTCCTACTATTTTCGATCAAGATGGTA 399
Db 3963 AAAAAATATATTAATAAAAAAACAATAAATACTATTATCTTTAAATATATCAAAATTTTTC 3904
QY 400 TTTTGTCTGTGAAGGAGAGAGTCTGACCCAGACCACTGCTTATTTCTTAATCAATCAT 459
Db 3903 CTTTCTCTATTTTTCATTAACCTTCTAACACAAAAACCACTACTATTATATATAAAAAA 3844
QY 460 ATGTTGTTATCTATCAAAAAAATAATATGATGATCATATATATGCTGGTTACTGTAA 515
Db 3843 AATAATTTTATTTTAACTAAACATAACTCATCTATATATCCCAAAACTTTTA 3788

RESULT 14
US-10-311-455-611
; Sequence 611, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 611
; LENGTH: 6650
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-611

Query Match      6.7%; Score 35.2; DB 15; Length 6650;
Best Local Similarity 68.1%; Pred. No. 51;
Matches 49; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 442 TTATTTCTAATCAATCATATGTTGTTATCTATCAAAAAAATAATGTCATCATGATATAT 501
Db 248 TTATTTTCTTATGATGTTTATGATGATGTTTATGATGATGATGATGATGATGATGATG 307
QY 502 GCTGTTACTGT 513
Db 308 TTTGGATGGGT 319

RESULT 15
US-10-311-455-352/c
; Sequence 352, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
```

Search completed: January 22, 2005, 07:25:42
Job time : 444 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 22, 2005, 04:53:35 ; Search time 83 Seconds
(without alignments)
4530.209 Million cell updates/sec

Title: US-10-725-829-1

Perfect score: 529
Sequence: 1 cataatggctgttcacaag.....ctgtaatggactttattcg 529

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	178.8	33.8	1360	3	US-08-454-295-2
2	178.8	33.8	1360	3	US-09-431-500A-2
3	178.8	33.8	1360	4	US-09-431-498-2
4	178.8	33.8	1360	4	US-09-431-499-2
5	149	28.2	1104	3	US-08-454-295-1
6	149	28.2	1104	3	US-09-431-500A-1
7	149	28.2	1104	4	US-09-431-498-1
8	149	28.2	1104	4	US-09-431-499-1
9	38.6	7.3	832	4	US-09-621-976-2813
10	35.8	6.8	1141	4	US-09-806-708B-22
11	34.8	6.6	6866	4	US-10-204-708-19
12	34.4	6.5	193303	4	US-09-497-855A-37
13	34.4	6.5	193303	4	US-09-497-855A-44
14	33.8	6.4	1141	4	US-09-806-708B-22
15	33.8	6.4	5473	4	US-09-620-312D-260
16	33.6	6.4	4043	4	US-09-710-279-3354
17	33.4	6.3	648	4	US-09-861-451A-25
18	33.2	6.3	2945	3	US-09-058-489-85
19	33.2	6.3	9439	3	US-09-058-489-89
20	33	6.2	777	4	US-09-540-236-1792
21	33	6.2	63563	4	US-09-596-002-33
22	32.8	6.2	12685	4	US-09-479-467A-3
23	32.6	6.2	558	4	US-09-134-000C-2786
24	32.6	6.2	1181	4	US-09-270-767-15127
25	32.4	6.1	832	4	US-09-621-976-2813
26	32	6.0	1314	4	US-09-248-796A-4870
27	31.8	6.0	1751	4	US-09-620-312D-847
					Sequence 2, Appl1
					Sequence 2, Appl1
					Sequence 2, Appl1
					Sequence 2, Appl1
					Sequence 1, Appl1
					Sequence 1, Appl1
					Sequence 1, Appl1
					Sequence 1, Appl1
					Sequence 2813, Ap
					Sequence 22, Appl
					Sequence 19, Appl
					Sequence 37, Appl
					Sequence 44, Appl
					Sequence 22, Appl
					Sequence 260, App
					Sequence 3354, Ap
					Sequence 25, Appl
					Sequence 85, Appl
					Sequence 89, Appl
					Sequence 1792, Ap
					Sequence 33, Appl
					Sequence 3, Appl1
					Sequence 2786, Ap
					Sequence 15127, A
					Sequence 2813, Ap
					Sequence 4870, Ap
					Sequence 847, App

Sequence 3, Appl1
Sequence 73, Appl1
Sequence 15, Appl
Sequence 2111, Ap
Sequence 10, Appl
Sequence 10, Appl
Sequence 1, Appl1
Sequence 5051, Ap
Sequence 29958, A
Sequence 584, App
Sequence 13891, A
Sequence 32, Appl
Sequence 25, Appl
Sequence 10, Appl
Sequence 33, Appl
Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-08-454-295-2

; Sequence 2, Application US/08454295

; Patent No. 6031087

; GENERAL INFORMATION:

; APPLICANT: Anderson, Marilyn A.

; APPLICANT: Atkinson, Angela H.

; APPLICANT: Heath, Robyn L.

; APPLICANT: Clarke, Adrienne E.

; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Scully, Scott, Murphy & Presser

; STREET: 400 Garden City Plaza

; CITY: Garden City

; STATE: New York

; COUNTRY: United States of America

; ZIP: 11530

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/454,295

; FILING DATE: 01-SEP-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Digiglio, Frank S.

; REGISTRATION NUMBER: 31,346

; REFERENCE/DOCKET NUMBER: 9748

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 742-4343

; TELEFAX: (516) 742-4366

; TELEX: 230 901 SANS UR

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1360 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 97..1200

; US-08-454-295-2

Query Match 33.8%; Score 178.8; DB 3; Length 1360;

Best Local Similarity 67.3%; Pred. No. 1.9e-46;

Matches 303; Conservative 0; Mismatches 132; Indels 15; Gaps 3;

US-09-431-498-2

Query Match 33.8%; Score 178.8; DB 4; Length 1360;
Best Local Similarity 67.3%; Pred. No. 1.9e-46;
Matches 303; Conservative 0; Mismatches 132; Indels 15; Gaps 3;

QY 5 ATGGCTGTTTCACAAAGTACCTTCCCTGCTGCTACTTGTCTTGGATGGATGTTCTA 64
DB 10 ATGGCTGTTTCACAGAGTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 69
QY 65 CT-----TGCAGAACATGTTGATGCAAGGCTTGTACTAGAGAAATGTGTCATT---TT 115
DB 70 GTAAGCAATGTGGAACATGCAATGCCAAGGCTTGTACCTTAACTGTGATCCAAAGTT 129
QY 116 AGCTATGGCATATGCCACGTTTCAGAGGAAGTCCCAAAACCTATATGACCAAAATGT 175
DB 130 GCCTATGGAGTTTGCCTGCTTTCAGAGAAAGAAAGAAATGATCGGATATGCAACACTGT 189
QY 176 TGCTCAGGCTATAGGGTTGCAACTATTACAGTCTTAAGGAGATTGATTGTGGAAGGA 235
DB 190 TGGCAGGCGACGAAGGTTGTAAGTACTTCAGTGATGATGGAACCTTTTGTGGAAGGA 249
QY 236 GAATCTGACCTAGAACCCCAAGGTTGCACTTATACAGTGTGTAAGGAGATTGTTGTAAGGA 295
DB 250 GAGTCTGATCTAGAAATCCAAAGGCTTGTACCTTAACTGTGATCCAAAGATTGCTAT 309
QY 296 TCAAAATGTCCTGTTTCAGAGGAAGATGATAATTAACCCCACTGGATGCAACACTTGT 355
DB 310 GGAGTTTGCCTGCTTTCAGAGGAAGAAAGAAAG-----AATGATCGGATATGCAACACTGT 363
QY 356 TGCAAGGCTATCAGGTTGCTACTATTTCGATCAAGATGGTGAATTTGTCTGTGAAGGA 415
DB 364 TGGCAGGCGACGAAGGTTGTAAGTACTTCAGTGATGATGGAACCTTTTGTGGAAGGA 423
QY 416 GAGAGTCTGAAACCCCAAGCCTGCTTAT 445
DB 424 GAGTCTGATCCTAGAAATCCAAAGGCTTGT 453

RESULT 4

US-09-431-499-2

; Sequence 2, Application US/09431499
; Patent No. 6451573
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,499
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGioglio, Frank S.
; REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 9748

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1360 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 97..1200

US-09-431-499-2

Query Match 33.8%; Score 178.8; DB 4; Length 1360;

Best Local Similarity 67.3%; Pred. No. 1.9e-46;

Matches 303; Conservative 0; Mismatches 132; Indels 15; Gaps 3;

QY 5 ATGGCTGTTTCACAAAGTACCTTCCCTGCTGCTACTTGTCTTGGATGGATGTTCTA 64
DB 10 ATGGCTGTTTCACAGAGTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 69
QY 65 CT-----TGCAGAACATGTTGATGCAAGGCTTGTACTAGAGAAATGTGTCATT---TT 115
DB 70 GTAAGCAATGTGGAACATGCAATGCCAAGGCTTGTACCTTAACTGTGATCCAAAGTT 129
QY 116 AGCTATGGCATATGCCACGTTTCAGAGGAAGTCCCAAAACCTATATGCAACAAATGT 175
DB 130 GCCTATGGAGTTTGCCTGCTTTCAGAGAAAGAAAGAAATGATCGGATATGCAACACTGT 189
QY 176 TGCTCAGGCTATAGGGTTGCACTTATACAGTGTGTAAGGAGATTGTTGTAAGGA 235
DB 190 TGGCAGGCGACGAAGGTTGTAAGTACTTCAGTGATGATGGAACCTTTTGTGTAAGGA 249
QY 236 GAATCTGACCTAGAACCCCAAGGTTGCACTTATACAGTGTGTAAGGAGATTGTTGTAAGGA 295
DB 250 GAGTCTGATCCTAGAAATCCAAAGGCTTGTACCTTAACTGTGATCCAAAGATTGCTAT 309
QY 296 TCAAAATGTCCTGTTTCAGAGGAAGATGATAATTAACCCCACTGGATGCAACACTTGT 355
DB 310 GGAGTTTGCCTGCTTTCAGAGGAAGAAAGAAAG-----AATGATCGGATATGCAACACTGT 363
QY 356 TGCAAGGCTATCAGGTTGCTACTATTTCGATCAAGATGGTGAATTTGTCTGTGAAGGA 415
DB 364 TGGCAGGCGACGAAGGTTGTAAGTACTTCAGTGATGATGGAACCTTTTGTGTAAGGA 423
QY 416 GAGAGTCTGAAACCCCAAGCCTGCTTAT 445
DB 424 GAGTCTGATCCTAGAAATCCAAAGGCTTGT 453

RESULT 5

US-08-454-295-1

; Sequence 1, Application US/08454295

; Patent No. 6031087

; GENERAL INFORMATION:

; APPLICANT: Anderson, Marilyn A.

; APPLICANT: Atkinson, Angela H.

; APPLICANT: Heath, Robyn L.

; APPLICANT: Clarke, Adrienne E.

; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Scully, Scott, Murphy & Presser

; STREET: 400 Garden City Plaza

; CITY: Garden City

; STATE: New York

; COUNTRY: United States of America

; ZIP: 11530

; COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1104 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-454-295-1

Query Match      28.2%; Score 149; DB 3; Length 1104;
Best Local Similarity 67.2%; Pred. No. 4.5e-37;
Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

QY 86 AAGGCTTGACTAGAGAAATGTGTC---ATTTAGCTATGGCATATGCCACGTTTCAGAA 142
Db 1 AAGGCTTGACTTAAACTGTGATCCAAAGAAATTCGCTATGGAGTTTGGCCGCGTTTCAGAA 60

QY 143 GGAAGTCCCCAAAACCTATATGCACCAATTTGTTGCTCAGGCTATTAAGGGTTTCAACTAT 202
Db 61 GAAAAGAAGAAATGATCGGATATGCACCAATTTGTTGCGCAGGCACGAAGGGTTTGAAGTAC 120

QY 203 TACAGTGTCTAAAGGAGATTTGATTTGTTGAAGGAGAAATCTGACCTAGAAACCCAAAAGAT 262
Db 121 TTCAATGATGATGGAACCTTTGTTGTTGAAGGAGAGTCTGATCTAGAAAATCCAAAGGCT 180

QY 263 TGTACCTTCGAATGTGATACACAGATTTGCTTATTCAAAATGTCTCTCGTTTCAGAAAGAAAG 322
Db 181 TGTACCTTAAACTGTGATCCAAAGAAATTCCTATGGAGTTTGGCCGCGTTTCAGAAAGAAAG 240

QY 323 ATGATAATTAACCCACTGGATGCACCACTTTGTTGACCGGCTATCAGGGTTGCTACTAT 382
Db 241 AAG-----AATGATCGGATATGCACCAACTTTGTTGCGCAGGCACGAAGGGTTTGAAGTAC 294

QY 383 TTCGATCAAGATGGTGTGATTTTGTCTGTGAAGGAGAGAGTCTCTGAACCCCAAGACCACTGCT 442
Db 295 TTCAGTGTATGGAACCTTTGTTGTTGAAGGAGAGTCTGATCTAGAAAATCCAAAGGCT 354

443 TAT 445
355 TGT 357

RESULT 6
US-09-431-500A-1
; Sequence 1, Application US/09431500A
; Patent No. 6261821
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,498
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/454,295
; CURRENT FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; ATTORNEY/AGENT INFORMATION:
```

```
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Nicotiana alata
US-09-431-500A-1

Query Match      28.2%; Score 149; DB 3; Length 1104;
Best Local Similarity 67.2%; Pred. No. 4.5e-37;
Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

QY 86 AAGGCTTGACTAGAGAAATGTGTC---ATTTAGCTATGGCATATGCCACGTTTCAGAA 142
Db 1 AAGGCTTGACTTAAACTGTGATCCAAAGAAATTCGCTATGGAGTTTGGCCGCGTTTCAGAA 60

QY 143 GGAAGTCCCCAAAACCTATATGCACCAATTTGTTGCTCAGGCTATTAAGGGTTTCAACTAT 202
Db 61 GAAAAGAAGAAATGATCGGATATGCACCAATTTGTTGCGCAGGCACGAAGGGTTTGAAGTAC 120

QY 203 TACAGTGTCTAAAGGAGATTTGATTTGTTGAAGGAGAAATCTGACCTAGAAACCCAAAAGAT 262
Db 121 TTCAATGATGATGGAACCTTTGTTGTTGAAGGAGAGTCTGATCTAGAAAATCCAAAGGCT 180

QY 263 TGTACCTTCGAATGTGATACACAGATTTGCTTATTCAAAATGTCTCTCGTTTCAGAAAGAAAG 322
Db 181 TGTACCTTAAACTGTGATCCAAAGAAATTCCTATGGAGTTTGGCCGCGTTTCAGAAAGAAAG 240

QY 323 ATGATAATTAACCCACTGGATGCACCACTTTGTTGACCGGCTATCAGGGTTGCTACTAT 382
Db 241 AAG-----AATGATCGGATATGCACCAACTTTGTTGCGCAGGCACGAAGGGTTTGAAGTAC 294

QY 383 TTCGATCAAGATGGTGTGATTTTGTCTGTGAAGGAGAGAGTCTCTGAACCCCAAGACCACTGCT 442
Db 295 TTCAGTGTATGGAACCTTTGTTGTTGAAGGAGAGTCTGATCTAGAAAATCCAAAGGCT 354

443 TAT 445
355 TGT 357

RESULT 7
US-09-431-498-1
; Sequence 1, Application US/09431498
; Patent No. 6440727
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,498
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/454,295
; CURRENT FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; ATTORNEY/AGENT INFORMATION:
```

```

; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1104 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-431-498-1

Query Match      28.2%; Score 149; DB 4; Length 1104;
Best Local Similarity 67.2%; Pred. No. 4.5e-37;
Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

QY 86 AAGGCTTGTAAGAGAAATGTGTC---ATTTAGCTATGGCATATGCCACGTTTCAGAA 142
DB 1 AAGGCTTGTAAGAGAAATGTGTC---ATTTAGCTATGGCATATGCCACGTTTCAGAA 60

QY 143 GGAAGTCCCAAAACCTATATGCACCAATTTGTCAGGCTATTAAGGTTTCAACTAT 202
DB 61 GAAAGAAGAATGATCGGATATGCACCACTGTTGCGCAGGCACGAAGGGTTGTAAGTAC 120

QY 203 TACAGTGTCTAAAGGAGATTTGATTTGTAAGGAGATCTGACCTAGAACCCCAAAAGAT 262
DB 121 TTCACTTAACTGTGATCAAGAAATTCCTATGGAGTTTGCCTATAGAAATCCAAAGGCT 180

QY 263 TGTACCTTGAATGTGATACAGATTTGCTTATTAATAATGTCCTCGTTTCAGAGGAAAG 322
DB 181 TGTACCTTAACTGTGATCAAGAAATTCCTATGGAGTTTGCCTATAGAAATCCAAAGGCT 240

QY 323 ATGATAATTAACCCACTGATGCACCACTTTGTCAGCGGCTATCAGGGTTGCTACTAT 382
DB 241 AAG-----AATGATCGGATATGCACCACTGTTGCGCAGGCACGAAGGGTTGTAAGTAC 294

QY 383 TTCGATCAAGATGCTGATTTTGTCTGTGAAGGAGAGAGTCTCTGAACCCAGACCACTGCT 442
DB 295 TTCAGTGTATGATGGAACCTTTTGTGTAAGGAGAGTCTGATCCTAGAAATCCAAAGGCT 354

QY 443 TAT 445
DB 355 TGT 357

RESULT 8
US-09-431-499-1
; Sequence 1, Application US/09431499
; Patent No. 6451573
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF
; TITLE OF INVENTION: AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,499
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1104 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-09-431-499-1

Query Match      28.2%; Score 149; DB 4; Length 1104;
Best Local Similarity 67.2%; Pred. No. 4.5e-37;
Matches 244; Conservative 0; Mismatches 110; Indels 9; Gaps 2;

QY 86 AAGGCTTGTAAGAGAAATGTGTC---ATTTAGCTATGGCATATGCCACGTTTCAGAA 142
DB 1 AAGGCTTGTAAGAGAAATGTGTC---ATTTAGCTATGGCATATGCCACGTTTCAGAA 60

QY 143 GGAAGTCCCAAAACCTATATGCACCAATTTGTCAGGCTATTAAGGTTTCAACTAT 202
DB 61 GAAAGAAGAATGATCGGATATGCACCACTGTTGCGCAGGCACGAAGGGTTGTAAGTAC 120

QY 203 TACAGTGTCTAAAGGAGATTTGATTTGTAAGGAGATCTGACCTAGAACCCCAAAAGAT 262
DB 121 TTCACTTAACTGTGATCAAGAAATTCCTATGGAGTTTGCCTATAGAAATCCAAAGGCT 180

QY 263 TGTACCTTGAATGTGATACAGATTTGCTTATTAATAATGTCCTCGTTTCAGAGGAAAG 322
DB 181 TGTACCTTAACTGTGATCAAGAAATTCCTATGGAGTTTGCCTATAGAAATCCAAAGGCT 240

QY 323 ATGATAATTAACCCACTGATGCACCACTTTGTCAGCGGCTATCAGGGTTGCTACTAT 382
DB 241 AAG-----AATGATCGGATATGCACCACTGTTGCGCAGGCACGAAGGGTTGTAAGTAC 294

QY 383 TTCGATCAAGATGCTGATTTTGTCTGTGAAGGAGAGAGTCTCTGAACCCAGACCACTGCT 442
DB 295 TTCAGTGTATGATGGAACCTTTTGTGTAAGGAGAGTCTGATCCTAGAAATCCAAAGGCT 354

QY 443 TAT 445
DB 355 TGT 357

RESULT 9
US-09-621-976-2813/c
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match
Best Local Similarity 14.1%; Pred. No. 0.028;
Matches 38; Conservative 116; Mismatches 115; Indels 0; Gaps 0;

QY 13 TCACAAAGTAGCTCTCTGCTTGCCTACCTGCTCTGCTGATGATGTTCTTACTTGGAA 72
Db 380 TAAATAATATATATTTGTYTWTWKKYTWYWTYTRMMWKKKARWYWMKSTYACAS 321

QY 73 ACATGTTGATGCCAAGCTGTACTAGAGAATGTGGTCAATTTAGCTATGGCATATGCC 132
Db 320 RYKYTWGWWYWMKMMSTRWTCYCKKCMYRGRCAYWTWARGMWSYANGKKSMR 261

QY 133 ACCTTCAGAGGAAGTCCCCAAAACCTATATGACCAATTTGTTGCTCAGGCTATAAGG 192
Db 260 SAMSMCTRMYYKGGSTVWTKTCATWCYWKYKRWMSKTCWSGSRGGYMTSYSTRS 201

QY 193 TTCCAACTATTACAGTCTAAAGAGAGATTGATTGTTGAAGGAGAATCTGACCCTAGAA 252
Db 200 YSMYASWMTWCMWGRWWSYTYWAWGKKWRYATTWRRAMWMAAWTMMWYMWNAW 141

QY 253 CCAAAAGATTGTACCTTCCAAATGTGATA 281
Db 140 CMSRGAAMYRRITMMWGHYWRKKSIR 112

RESULT 10
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEL promoters
US-09-806-708B-22

Query Match
Best Local Similarity 11.5%; Pred. No. 0.26;
Matches 44; Conservative 159; Mismatches 180; Indels 2; Gaps 1;

QY 96 CTAGAGAATGTGCTATTTTAGCTATGGCATATGCCACGTTTCAGAGGAAGTCCCAAA 155
Db 581 YTRNTYCKSYAHSYWWSNNMYRRYSARNSSMARWTTNNWWSGBVRMRWAGTWM 522

QY 156 AACCTATATGCACCAATTTGTTGCTAGGCTATAAGGGTTGCAACTATTACAGTCTAAG 215
Db 521 WRHNNNTDTRYWYWKWREBTTTYDSMCNAKSMWRGNWNRWAKMWWAANNADAGMD 462

QY 216 GAGATTGTTGTTGAAGGAGAATCTGACCTAGAACCCAAAGAGATTGTACCTCGAAT 275
Db 461 HTYWMGNNTMMWRWAKMNNWACWRRAYCCNNNNRACVWHKHWRWTKWTKWKAACN 402

QY 276 GTGATACACAGATTGCTTATTTCAAAATGTCTCTGTTTCAAGGAAGAGATGATAATTAAC 335
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match
Best Local Similarity 14.1%; Pred. No. 0.028;
Matches 38; Conservative 116; Mismatches 115; Indels 0; Gaps 0;

QY 13 TCACAAAGTAGCTCTCTGCTTGCCTACCTGCTCTGCTGATGATGTTCTTACTTGGAA 72
Db 380 TAAATAATATATATTTGTYTWTWKKYTWYWTYTRMMWKKKARWYWMKSTYACAS 321

QY 73 ACATGTTGATGCCAAGCTGTACTAGAGAATGTGGTCAATTTAGCTATGGCATATGCC 132
Db 320 RYKYTWGWWYWMKMMSTRWTCYCKKCMYRGRCAYWTWARGMWSYANGKKSMR 261

QY 133 ACCTTCAGAGGAAGTCCCCAAAACCTATATGACCAATTTGTTGCTCAGGCTATAAGG 192
Db 260 SAMSMCTRMYYKGGSTVWTKTCATWCYWKYKRWMSKTCWSGSRGGYMTSYSTRS 201

QY 193 TTCCAACTATTACAGTCTAAAGAGAGATTGATTGTTGAAGGAGAATCTGACCCTAGAA 252
Db 200 YSMYASWMTWCMWGRWWSYTYWAWGKKWRYATTWRRAMWMAAWTMMWYMWNAW 141

QY 253 CCAAAAGATTGTACCTTCCAAATGTGATA 281
Db 140 CMSRGAAMYRRITMMWGHYWRKKSIR 112

RESULT 11
US-10-204-708-19
; Sequence 19, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 19
; LENGTH: 6866
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-19

Query Match
Best Local Similarity 52.0%; Pred. No. 1.4;
Matches 78; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 380 TATTTCGATCAAGATGCTGATTTTGTCTGTGAGGAGAGAGTCTGACCCAGACCACT 439
Db 1648 TAGTTGGCGGATGAGGTGAGATTTTGTGTTTAAATAAAAAAAAAAAAAAAAAAAT 1707

QY 440 GCTTATTCTTAATCAATCATATGTTCTATCATCAAAAAAAAAAATATGATGATGAT 499
Db 1708 TTTTATTAGATTGAAGTTTATTATTAGGNTAATAAATAATGTTAGTTATATTT 1767

QY 500 ATGCTGGTTACTGTAATGTGGACTTTATTG 529
Db 1768 GTTATAATTTATAAAAGTTTATTTTTTTAG 1797

RESULT 12
US-09-497-855A-37
; Sequence 37, Application US/09497855A
; Patent No. 6605432
; GENERAL INFORMATION:
; APPLICANT: Huang, Tim
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
; FILE REFERENCE: UMO1523
```

; CURRENT APPLICATION NUMBER: US/09/497,855A
; CURRENT FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/120,592
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 60/118,760
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 193303
; TYPE: DNA
; ORGANISM: Homo sapiens;
US-09-497-855A-37

Query Match 6.5%; Score 34.4; DB 4; Length 193303;
Best Local Similarity 52.9%; Pred. No. 11;
Matches 74; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 218 GATTTCATTGTCAGGAGATCTGACCCCTAGAAACCCAAAGATTGTACCTTCGAATGT 173233
DB 173174 GACTTCAGAGCTATTGGAGAAATCAGAGGGTAAATTCACAGAAATCTCTAAATTTAAATGT 173233
QY 278 GATACACAGATTCCTTATTCAAAATGTCCTCGTTTCAGAGGAAGATGATAATTAACCC 337
DB 173234 GACTAAAGTTTATATCTTCTTCAGTTCTGCTCATGAAGTTGAGAGAAATGAATAATGCG 173293
QY 338 ACTGGATGCACCACTTGTTG 357
DB 173294 ATTCACGGAGCATTTGTTG 173313

RESULT 13
US-09-497-855A-44
; Sequence 44, Application US/09497855A
; Patent No. 6605432
; GENERAL INFORMATION:
; APPLICANT: Huang, Tim
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
; FILE REFERENCE: UMO1523
; CURRENT APPLICATION NUMBER: US/09/497,855A
; CURRENT FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/120,592
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 60/118,760
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 193303
; TYPE: DNA
; ORGANISM: Homo sapiens;
US-09-497-855A-44

Query Match 6.5%; Score 34.4; DB 4; Length 193303;
Best Local Similarity 52.9%; Pred. No. 11;
Matches 74; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 218 GATTTCATTGTCAGGAGATCTGACCCCTAGAAACCCAAAGATTGTACCTTCGAATGT 173233
DB 173174 GACTTCAGAGCTATTGGAGAAATCAGAGGGTAAATTCACAGAAATCTCTAAATTTAAATGT 173233
QY 278 GATACACAGATTCCTTATTCAAAATGTCCTCGTTTCAGAGGAAGATGATAATTAACCC 337
DB 173234 GACTAAAGTTTATATCTTCTTCAGTTCTGCTCATGAAGTTGAGAGAAATGAATAATGCG 173293
QY 338 ACTGGATGCACCACTTGTTG 357
DB 173294 ATTCACGGAGCATTTGTTG 173313

RESULT 14
US-09-806-708B-22
; Sequence 22, Application US/09806708B

; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: Consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

Query Match 6.4%; Score 33.8; DB 4; Length 1141;
Best Local Similarity 11.1%; Pred. No. 1.1;
Matches 40; Conservative 136; Mismatches 185; Indels 0; Gaps 0;
QY 131 CCAGGTTCAAGAGGAAGTCCCAAAACCTATATGACCAATTTGTTGCTCAGGCTATAAG 190
DB 345 VYNNNNNNNTYKKARHBARDWVHSAWKWHANAHAHYSRKKWTBYKRKTVMVNNNGT 404
QY 191 GGTTCGAACCTATTACAGTCTAAAGGAGATTGATTTGTTGAAGGAGAGACTGACCCCTAGA 250
DB 405 TWMKRMWAWYKMDMDMBGTNNNNNGRTYTGTTKNNKMMWYKWKANNCKRWDHKT 464
QY 251 AACCCAAAAGATTGACCTTCGAATGTGATACACAGATTGCTTTATTCAAATGTCCTCGT 310
DB 465 CTHNNTTWMKMTYNNNCYKSWTNGSKHRBAAAVTYWMMWRRYAHANNNDYWKKA 524
QY 311 TCAGAGGAAAGATGATAATTAACCCCTGAGATGACCACTTTGTTGACGGCTATCAG 370
DB 525 CTWYKYBVCCKWNNYAAYTKSSWNTSRYYRWKTNNSWRSDTRSMGRANNYARABH 584
QY 371 GGTTCGCTACTATTTCGATCAAGATGGTGATTTGTTGTTGAAGGAGAGAGTCTCGAACCC 430
DB 585 YGYKMNTRWBSHTWBHBRAGAAHYMMBMVBAKCHWKAWYKAKKYAGAGGSSNNNNN 644
QY 431 AAGACCCTGCTATTCTTAATCAATCATATGTTGTTATCTATCAAAAAAATATGTAT 490
DB 645 NNNNNNNNATCARDYYAASRWYANAKWYKYKBAANNAYYTHANNWGCWNNATDT 704
QY 491 G 491
DB 705 R 705

RESULT 15
US-09-620-312D-260
; Sequence 260, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenchua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui

; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 656962el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 260
; LENGTH: 5473
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4482)
US-09-620-312D-260

Query Match 6.4%; Score 33.8; DB 4; Length 5473;
Best Local Similarity 48.7%; Pred. No. 2.5;
Matches 92; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
Qy 125 ATATGCCCAGTTTCAGAGGAAGTCCCAAAACCTATATGCACCAATTGTTGCTCAGGC 184
Db 4988 ATATGATCACATGTAGAAGTAGCTTCAAGAAATTTCTTGTTAGTCATAAATGTTTAAATAA 5047
Qy 185 TATAAGGGTTGCCAACTATTACAGTGTCTAAAGGAGATTGATTTGTGAAGGAGAACTCTGAC 244
Db 5048 TATGATGTAAATTTATTTATGAGCCTATGATGATGATGATGATGATGATGATGATGATGAT 5107
Qy 245 CCTAGAAACCCAAAGATTCTACCTTCGAATGTGATACACAGATTGCTTATTCAAAATGT 304
Db 5108 CTTAGAAAATGTAGACATGTTTAACTGGGAAATAAAATATAGAGTGGCACTTCAAGACAA 5167
Qy 305 CCTCGTTCA 313
Db 5168 GCTGACTCA 5176

Search completed: January 22, 2005, 07:18:15
Job time : 88 secs

proteinase inhibitor II

Db	1	MAMHKEVNFAYLLIVLGNFL---YDAKACTRECNLGFICPRSEGSPLNPICINCCS	57
Qy	60	GKYGKNTYSARAGDLICEGESDPRNPKDCTFECDQTIAYSKCPRSEGKMIIKPTGCTTCCT	119
Db	58	GKYGKNTYSNGFKFICEGESDPKRPNACFTNCDFPNIAYSRCPRSOGKSLIYPTGCTTCT	117
Qy	120	GYGCGYFPDQDGDFVCCEGSPEPKTTAY	147
Db	118	GKYGCTYFGDKGKFVCEGESDEPKANMY	145
 RESULT 4 T07597			
proteinase inhibitor II - potato			
C/Species: Solanum tuberosum (potato)			
C/Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004			
C/Accession: T07597			
R/Lee, J.S.			
submitted to the EMBL Data Library, June 1992			
A/Reference number: S24965			
A/Accession: T07597			
A/Status: preliminary; translated from GB/EMBL/DDBJ			
A/Molecule type: DNA			
A/Residues: 1-158 <LBE>			
A/Cross-references: UNIPROT.Q00782; EMBL.Z12753; NID:g21553; PIDN:CAA78277.1; PID:g21554			
A/Experimental source: cv. Russet Burbank			
C/Genetics:			
A/Introns: 17/1			
C/Superfamily: potato proteinase inhibitor PTI			
C/Keywords: serine proteinase inhibitor			
 Query Match 69.4%; Score 593.5; DB 2; Length 158; Best Local Similarity 71.9%; Pred. No. 1.3e-41; Matches 105; Conservative 12; Mismatches 26; Indels 3; Gaps 2;			
Qy	1	MAMHK-VSFACLILVLGMFLLA--KHVDAKATRECCHFSYGCIPRSESGSPQPICTNC	57
Db	1	MAIHKEVSFLAYLLVLGLMLLFVSAMEHVDAKACTLECNLGYGICPRSEGSPEPICITNC	60
Qy	58	CSGYKGNYSAKGDLICEGESDPRNPKDCTFECDQTIAYSKCPRSEGKMIIKPTGCTTC	117
Db	61	CAGYKGNYISANGTFICEGSHPKNPKACPKNCDPHIAYSRCPRSGGKTLIYPTGCTTC	120
Qy	118	CTGYGCGYFPDQDGDFVCCEGSPEPK	143
Db	121	CTGYTDICYFGDKGKFVCEGESIEPK	146
 RESULT 5 S24973			
proteinase inhibitor II - potato			
C/Species: Solanum tuberosum (potato)			
C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004			
C/Accession: S24973			
R/Choi, Y.; Kim, J.W.; Lee, J.S.			
submitted to the EMBL Data Library, July 1992			
A/Description: Characterization of a potato proteinase inhibitor II gene that is expressed			
A/Reference number: S24973			
A/Accession: S24973			
A/Status: preliminary			
A/Molecule type: DNA			
A/Residues: 1-147 <CHO>			
A/Cross-references: UNIPROT.Q41489; EMBL.Z13992; NID:g21555; PIDN:CAA78383.1; PID:g21556			
C/Genetics:			
A/Introns: 18/1			
C/Superfamily: potato proteinase inhibitor PTI			
 Query Match 68.5%; Score 586; DB 2; Length 147; Best Local Similarity 72.2%; Pred. No. 5.1e-41; Matches 104; Conservative 12; Mismatches 24; Indels 4; Gaps 2;			
Qy	1	MAMHK-VSFACLILVLGMFLLA-KHVDAKATRECCHFSYGCIPRSESGSPQPICTNC	59

Db 1 MAVHKEVSVAYLLVLVGMFL---YVDALCTKEGNGLFGICPRSESGSPNPICTNC 57
QY 60 GYKGCNYSAGDLICGESDPRNPDKCTPECDTQIAYSKCPRSESGMKMIKPTGCTTC 119
Db 58 GYKGCNYSAGDLICGESDPRNPDKCTPECDTQIAYSKCPRSESGMKMIKPTGCTTC 117
QY 120 GYKGCNYSAGDLICGESDPRNPDKCTPECDTQIAYSKCPRSESGMKMIKPTGCTTC 143
Db 118 GYKGCNYSAGDLICGESDPRNPDKCTPECDTQIAYSKCPRSESGMKMIKPTGCTTC 141

RESULT 6
XKROCI
proteinase inhibitor PCI-I precursor - potato
N;Alternate names: potato chymotrypsin inhibitor I; protease inhibitor II
C;Species: Solanum tuberosum (potato)
C;Date: 06-Jul-1982 #sequence revision 12-Apr-1996 #text_change 09-Jul-2004
C;Accession: A26584; A01319; A23591
R;Thornburg, R.W.; An, G.; Cleveland, T.E.; Johnson, R.; Ryan, C.A.
Proc. Natl. Acad. Sci. U.S.A. 84, 744-748, 1987
A;Title: Wound-inducible expression of a potato inhibitor II-chloramphenicol acetyltransferase
A;Reference number: A26584
A;Accession: A26584
A;Molecule type: DNA
A;Residues: 1-153 <THO>
A;Cross-references: UNIPROT:P01080; GB:M15186; NID:g169488; PIDN:AAA33815.1; PID:g169489
R;Hass, G.M.; Hermanson, M.A.; Ryan, C.A.; Gentry, L.
Biochemistry 21, 752-756, 1982
A;Title: Primary structures of two low molecular weight proteinase inhibitors from potato
A;Reference number: A90465; MUID:82182863; PMID:7074039
A;Accession: A01319
A;Molecule type: protein
A;Residues: 55-106 <HAS>
A;Note: Leu-92 is probably the site of interaction with chymotrypsin
R;Keil, M.; Sanchez-Serrano, J.; Schell, J.; Willmitzer, L.
Nucleic Acids Res. 14, 5641-5650, 1986
A;Title: Primary structure of a proteinase inhibitor II gene from potato (Solanum tuberosum)
A;Reference number: A23591; MUID:86286579; PMID:3016659
A;Accession: A23591
A;Molecule type: DNA
A;Residues: 1-26, 'EH', 28-33, 'TL', 36-54, 'R', 56-153 <KEI>
A;Cross-references: GB:X04118; NID:g21521; PIDN:CAA27730.1; PID:g21522
C;Genetics:
A;Gene: IIK
A;Introns: 18/1
A;Superfamily: potato proteinase inhibitor PTI
C;Keywords: serine proteinase inhibitor
F;1-25/Domain: signal sequence #status predicted <SIG>
F;55-106/Product: proteinase inhibitor PCI-I #status experimental <MAT>

Query Match 68.5%; Score 585.5; DB 1; Length 153;
Best Local Similarity 70.0%; Pred. No. 5.7e-41;
Matches 105; Conservative 14; Mismatches 28; Indels 3; Gaps 3;

QY 1 MAVHK-VSLFACL-LVLGMFL--AKHVDKACTREGHFGICPRSESGSPKPICTNC 57
Db 1 MDVHKEVNFVAYLLVLVGLVLSVMDVDKAKCTREGNGLFGICPRSESGSPNPICTNC 60
QY 58 CSYKGCNYSAGDLICGESDPRNPDKCTPECDTQIAYSKCPRSESGMKMIKPTGCTTC 117
Db 61 CAGYKGCNYSANGAFICEQSDPKPKKACPLNCDPHIAYSKCPRSESGSLIYPTGCTTC 120
QY 118 CTGYGCGYFDDGDGFVCEGESPEPKTAY 147
Db 121 CTGYKGCYFGKNGKFCVCEGESDEPKANMY 150

RESULT 7
S43105
proteinase inhibitor II - potato
C;Species: Solanum tuberosum (potato)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C;Accession: S43105
R;Murray, C.; Christeller, J.T.
submitted to the EMBL Data Library, March 1994
A;Reference number: S43105
A;Accession: S43105
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-154 <MUR>
A;Cross-references: UNIPROT:Q43652; EMBL:X78275; NID:g467609; PIDN:CAA55082.1; PID:g46761
C;Genetics:
A;Introns: 18/1
C;Superfamily: potato proteinase inhibitor PTI

Query Match 68.0%; Score 581; DB 2; Length 154;
Best Local Similarity 67.5%; Pred. No. 1.3e-40;
Matches 102; Conservative 16; Mismatches 29; Indels 4; Gaps 2;

QY 1 MAVHK-VSLFACL-LVLGMFL--AKHVDKACTREGHFGICPRSESGSPKPICTNC 56
Db 1 MDVHKEVNFVAYLLVLVGLVLSVMDVDKAKCTREGNGLFGICPRSESGSPNPICTNC 60
QY 57 CSYKGCNYSAGDLICGESDPRNPDKCTPECDTQIAYSKCPRSESGMKMIKPTGCTTC 116
Db 61 CSYKGCNYSVFGRFICEGESDLAKNPKACPLNCDTNIAYSCPHSEGSLSIYPTGCTTC 120
QY 117 CTGYGCGYFDDGDGFVCEGESPEPKTAY 147
Db 121 CTGYKGCYFGKNGKFCVCEGESDEPKANMY 151

RESULT 8
JQ2153
proteinase inhibitor II precursor - Persian tobacco
C;Species: Nicotiana glauca (Persian tobacco)
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C;Accession: JQ2153; PQ0647; S65396
R;Atkinson, A.H.; Heath, R.L.; Simpson, R.J.; Clarke, A.E.; Anderson, M.A.
Plant Cell 5, 203-213, 1993
A;Title: Proteinase inhibitors in Nicotiana glauca stigmas are derived from a precursor p
A;Reference number: JQ2153; MUID:93200805; PMID:8453302
A;Accession: JQ2153
A;Molecule type: mRNA
A;Residues: 1-397 <ATK>
A;Cross-references: UNIPROT:Q40378; GB:U08219; NID:g473590; PIDN:AAA17739.1; PID:g473591
A;Accession: PQ0647
A;Molecule type: protein
A;Residues: 54-66; 112-124; 170-182; 228-240; 286-298; 344-357 <AT2>
A;Experimental source: stigma, style
R;Heath, R.L.; Barton, P.A.; Simpson, R.J.; Reid, G.E.; Lim, G.; Anderson, M.A.
Eur. J. Biochem. 230, 250-257, 1995
A;Title: Characterization of the protease processing sites in a multidomain proteinase p
A;Reference number: S65396; MUID:95324532; PMID:7601108
A;Accession: S65396
A;Molecule type: protein
A;Residues: 54-227; 286-343 <HEA>
C;Comment: This protein contains six similar domains, each with a potential active site.
C;Superfamily: potato proteinase inhibitor PTI
C;Keywords: serine proteinase inhibitor
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-397/Product: proteinase inhibitor II #status predicted <MAT>
F;35-36, 93-94/Region: chymotrypsin-specific sites
F;153-154, 211-212, 269-270, 327-328/Region: trypsin-specific sites

Query Match 52.6%; Score 450; DB 2; Length 397;
Best Local Similarity 57.5%; Pred. No. 1.3e-29;
Matches 84; Conservative 14; Mismatches 42; Indels 6; Gaps 4;

QY 1 MAVHKEVSVAYLLVLVGMFL--AKHVDKACTREGHFGICPRSESGSPKPICTNC 57
Db 1 MDVHKEVSVAYLLVLVGLVLSVMDVDKAKCTREGNGLFGICPRSESGSPNPICTNC 60
QY 58 CSYKGCNYSAGDLICGESDPRNPDKCTPECDTQIAYSKCPRSESGMKMIKPTGCTTC 117

QY 53 ICTNCCSGYKGNYSYSAKGLDICEGSDPRNPXDKCTFECDTQIAYSKCPR 102
Db 2 ICTNCCAGYKGCNYYSANGAFICEGSDPKPNVCPNRCNDTNIAYSKCLR 51

RESULT 13

QY2269
trypsin inhibitor-1 - common tobacco
N;Alternate names: proteinase inhibitor
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 09-Jul-2004
C;Accession: JQ2269
R;Pearce, G.; Johnson, S.; Ryan, C.A.
Plant Physiol. 102, 639-644, 1993
A;Title: Purification and characterization from tobacco (Nicotiana tabacum) leaves of si
A;Reference number: JQ2269; MUID:94151442; PMID:8108514
A;Accession: JQ2269
A;Molecule type: protein
A;Residues: 1-53 <PEA>
A;Cross-references: UNIPROT:Q7M1P5
A;Experimental source: leaf
A;Superfamily: potato proteinase inhibitor PTI
C;Keywords: serine proteinase inhibitor

Query Match 23.3%; Score 199; DB 2; Length 53;
Best Local Similarity 64.7%; Pred. No. 8e-10;
Matches 33; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSYSAKGLDICEGSDPRNPXDKCTFECDTQIAYSKCPRS 103
Db 3 ICTNCCAGYKGCYFYSDDTFTVCEGSDPRNPXACPRNCDPRIAYGICPLS 53

RESULT 14

TIB01
proteinase inhibitor - eggplant
C;Species: Solanum melongena (eggplant, aubergine)
C;Date: 30-Nov-1980 #sequence_revision 30-Nov-1980 #text_change 09-Jul-2004
C;Accession: A01317
R;Richardson, M.
FEBS Lett. 104, 322-326, 1979
A;Title: The complete amino acid sequence and the trypsin reactive (inhibitory) site of
A;Reference number: A01317; MUID:80004163; PMID:477995
A;Accession: A01317
A;Molecule type: protein
A;Residues: 1-52 <RIC>
A;Cross-references: UNIPROT:P01078
A;Note: 2-Leu and 6-Cys were also found
C;Superfamily: potato proteinase inhibitor PTI
C;Keywords: pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;38/Inhibitory site: Arg (trypsin) #status experimental

Query Match 20.5%; Score 175; DB 1; Length 52;
Best Local Similarity 58.8%; Pred. No. 7.1e-08;
Matches 30; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSYSAKGLDICEGSDPRNPXDKCTFECDTQIAYSKCPRS 103
Db 2 ICTNCCAGYKGCYFYSDDTFTVCEGSDPRNPXACPRNCDGRIAYGICPLS 52

RESULT 15

XKPO2A
proteinase inhibitor IIa - potato (fragment)
C;Species: Solanum tuberosum (potato)
C;Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 09-Jul-2004
C;Accession: A01320
R;Iwasaki, T.; Kiyohara, T.; Yoshikawa, M.
J. Biochem. 79, 381-391, 1976
A;Title: Amino acid sequence of an active fragment of potato proteinase inhibitor IIa.
A;Reference number: A01320; MUID:76190064; PMID:1270410
A;Accession: A01320

A;Molecule type: protein
A;Residues: 1-45 <IWA>

A;Cross-references: UNIPROT:P01081
A;Note: this active fragment inhibits trypsin strongly and chymotrypsin temporarily; the
C;Superfamily: potato proteinase inhibitor PTI
C;Keywords: serine proteinase inhibitor
F;10-24, 14-35, 20-43/Disulfide bonds: #status experimental
F;32/Inhibitory site: Lys (trypsin) #status experimental

Query Match 14.9%; Score 127; DB 1; Length 45;
Best Local Similarity 57.1%; Pred. No. 0.00052;
Matches 24; Conservative 4; Mismatches 10; Indels 4; Gaps 1;

QY 45 SEGSPQKPICTNCCSGYKGCNYYS----AKGDLICEGSDPR 82
Db 1 SEGSPENRICTNCCAGYKGCNYNCNTNIAYSKVCEGFDPK 42

Search completed: January 24, 2005, 16:52:48
Job time : 38 secs

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2005, 16:43:10 ; Search time 192 Seconds

(without alignments)
443.518 Million cell updates/sec

Title: US-10-725-829-2

Perfect score: 855

Sequence: 1 MAVHKVSFLACLLVLGMFL.....QDGFVCEGESPEPKTTAYF 148

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	855	100.0	148	2 O8W2K6	O8W2K6 solanum ame
2	614.5	71.9	197	1 IP21_TOBAC	Q40561 nicotiana t
3	612	71.6	152	2 O6JKN9	O6JKN9 solanum nig
4	612	71.6	152	2 AAR37362	AAR37362 solanum n
5	602	70.4	201	1 IP23_LYCES	Q43502 lycopersico
6	596	69.7	148	1 IP21_LYCES	P05119 lycopersico
7	596	69.7	148	2 AAG12170	AAG12170 lycopersi
8	596	69.7	148	2 BAC76901	BAC76901 lycopersi
9	596	69.7	148	2 AAN05016	AAN05016 lycopersi
10	595	69.6	147	1 IP2T_SOLTU	Q41435 solanum tub
11	595	69.6	154	2 O84L56	Q84L56 solanum phu
12	595	69.6	154	2 AAS01731	AAS01731 solanum p
13	593.5	69.4	158	1 IP2X_SOLTU	Q00782 solanum tub
14	589	68.9	154	1 IP25_SOLTU	Q41488 solanum tub
15	586	68.5	147	1 IP2Y_SOLTU	Q41489 solanum tub
16	585.5	68.5	153	1 IP2K_SOLTU	P01080 solanum tub
17	581	68.0	154	1 IP27_SOLTU	Q43652 solanum tub
18	555	64.9	126	2 O82735	O82735 solanum tub
19	487	57.0	109	2 O8W2K1	Q8W2K1 solanum ame
20	450	52.6	397	2 Q40378	Q40378 nicotiana a
21	449	52.5	397	2 Q9S077	Q9S077 nicotiana a
22	440	51.5	455	2 O84L39	O84L39 nicotiana a
23	430	50.3	204	1 IP22_CAPAN	Q49146 capsicum an
24	425.5	49.8	506	2 Q9SDW8	Q9SDW8 nicotiana g
25	425	49.7	204	2 Q9SDU4	Q9SDU4 capsicum an
26	416	48.7	390	2 Q9SDW7	Q9SDW7 nicotiana g
27	343.5	40.2	410	2 O6H238	O6H238 nicotiana a
28	334	39.1	146	2 Q7M222	Q7M222 lycopersico
29	333	38.9	223	1 IP22_LYCES	Q43710 lycopersico
30	245	28.7	55	1 IP21_CAPAN	P56615 capsicum an
31	231.5	27.1	143	2 Q9M514	Q9M514 capsicum an

32	226	26.4	51	1 IP21_SOLTU	P01079 solanum tub
33	199	23.3	53	2 Q7M1P5	Q7M1P5 nicotiana t
34	187	21.9	52	1 IPR_SOLME	P01078 solanum mel
35	180	21.1	58	2 O84LQ9	O84LQ9 nicotiana a
36	127	14.9	45	1 IP2A_SOLTU	P01081 solanum tub
37	121	14.2	52	1 IP23_CAPAN	P83241 capsicum an
38	111	13.0	40	1 IP2B_SOLTU	P01082 solanum tub
39	110	12.9	46	2 Q94I92	Q94I92 atropa bell
40	106	12.4	137	1 THN5_HORVU	P09617 hordeum vul
41	106	12.4	137	1 THN6_HORVU	Q8H0Q5 hordeum vul
42	105	12.3	1101	2 Q964D2	Q964D2 entamoeba h
43	103	12.0	320	2 O57079	O57079 cowpox viru
44	101.5	11.9	114	2 Q9S9A0	Q9S9A0 viscum albu
45	101	11.8	137	1 THN7_HORVU	Q42838 hordeum vul

ALIGNMENTS

RESULT 1

ID	Q8W2K6	PRELIMINARY;	PRT;	148 AA.
AC	Q8W2K6;			
DT	01-MAR-2002 (Tremblrel. 20, Created)			
DT	01-MAR-2002 (Tremblrel. 20, Last sequence update)			
DT	01-JUN-2003 (Tremblrel. 24, Last annotation update)			
DE	Proteinase inhibitor IIA.			
GN	Names=PIN2a;			
OS	Solanum americanum.			
OC	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;			
OC	lamids; Solanales; Solanaceae; Solanum.			
OX	NCBI_TaxID=109975;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21643209; PubMed=11785934;			
RA	Xu Z.-F., Qi W.-Q., Ouyang X.-Z., Yeung E., Chye M.-L.;			
RT	"A proteinase inhibitor II of Solanum americanum is expressed in			
RT	phloem.";			
RL	Plant Mol. Biol. 47:727-738 (2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Chye M.-L., Xu Z.-F.;			
RL	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF174381; AAL36458.1; --			
DR	HSSP; P05119; IOYV.			
DR	GO; GO:0004867; F:serine-type endopeptidase inhibitor activity; IEA.			
DR	InterPro; IPR003465; Prot inh Pin2.			
DR	Pfam; PF02428; Prot_inhib II_2.			
SQ	SEQUENCE 148 AA, 16323 MW, 053536378E104E75 CRC64;			

Query Match 100.0%; Score 855; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.1e-69;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAVHKVSFLACLLVLGMFLAKHVDKACTRECGHFSYGICPRSGSPKPICTNCCSG	60
Db	1	MAVHKVSFLACLLVLGMFLAKHVDKACTRECGHFSYGICPRSGSPKPICTNCCSG	60
Qy	61	YKGCNYSAGDILICGESDPRNPDKCTFECDTQIAYSKCPRSEGKMIKPTGCTTCGTG	120
Db	61	YKGCNYSAGDILICGESDPRNPDKCTFECDTQIAYSKCPRSEGKMIKPTGCTTCGTG	120
Qy	121	YQGCYFFDQDGFVCEGESPEPKTTAYF	148
Db	121	YQGCYFFDQDGFVCEGESPEPKTTAYF	148

RESULT 2

IP21_TOBAC				
ID	IP21_TOBAC	STANDARD;	PRT;	197 AA.
AC	Q40561;			
DT	13-DEC-1998 (Rel. 37, Created)			

```
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Proteinase inhibitor type II precursor.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Sameun NN; TISSUE=Leaf;
RX MEDLINE=95284369; PubMed=776901;
RA Balandin M.T., van der Does C., Albert J.M., Bol J.F.,
RA Linthorst H.J.M.;
RT "Structure and induction pattern of a novel proteinase inhibitor class
RT II gene of tobacco.";
RL Plant Mol. Biol. 27:1197-1204(1995).
CC -I- INDUCTION: Locally induced in leaves subjected to different types
CC of stress (TMV infection, wounding, UV irradiation).
CC -I- SIMILARITY: Belongs to the potato type II proteinase inhibitor
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z29537; CAA82652.1; -.
DR PIR; S56662; S56662.
DR HSP; P05119; IOYV.
DR InterPro; IPR003465; Prot_inh_Pin2.
DR Pfam; PF02428; Prot_inhib_II_3.
KW Repeat; Serine protease inhibitor; Signal.
FT SIGNAL 1 24 Potential.
FT CHAIN 25 197 Proteinase inhibitor type II.
FT REPEAT 24 80 1.
FT REPEAT 81 140 2.
FT REPEAT 141 196 3.
FT SITE 29 30 Reactive bond for trypsin (Potential).
FT DISULFID 27 115 By similarity.
FT DISULFID 31 111 By similarity.
FT DISULFID 39 121 By similarity.
FT DISULFID 51 88 By similarity.
FT DISULFID 54 72 By similarity.
FT DISULFID 55 84 By similarity.
FT DISULFID 61 97 By similarity.
FT DISULFID 114 132 By similarity.
FT SEQUENCE 197 AA; F9CD5CB4267A9710 CRC64;
Query Match 71.9%; Score 614.5; DB 1; Length 197;
Best Local Similarity 75.7%; Pred. No. 8.2e-48;
Matches 109; Conservative 12; Mismatches 20; Indels 3; Gaps 2;
QY 1 MAVHKVSFLACLVLGWMFLAKHVDKACTREGHFSYGCIPRSEGSPOKPICTNCCSG 60
DB 1 MAVHKVSFVAHLVLVG-MFLLL--VDKAKTCKGKNFAYGICPRSQGTDPDPICTTCCAG 57
QY 61 YKGCNYSYSAKGLICEGSDPRNPKDCTFECDTQIAYSKCPRSEGRKMIKPTGCTTCCGTG 120
DB 58 YKGCNYSYANGTIFCEGSSDPKPNVCPQFCDDPDIAYSKCPREGEETIINPTGCTTCCGTG 117
QY 121 YQCYTFDQDGFVCEGSPEPKT 144
DB 118 YKGCYTFGQDGFVCEGSDEPKS 141
RESULT 3
Q6JKN9 PRELIMINARY; PRT; 152 AA.
ID Q6JKN9
AC Q6JKN9;
```

```
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Proteinase inhibitor 2b precursor.
GN Name=PIN2b;
OS Solanum nigrum (Black nightshade).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4112;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15078438;
RA Schmidt D.D., Kessler A., Kessler D., Schmidt S., Lim M., Gase K.,
RA Baldwin I.T.;
RT "Solanum nigrum: a model ecological expression system and its tools.";
RL Mol. Ecol. 13:981-995(2004).
DR EMBL; AY422686; AAR37362.1; -.
DR InterPro; IPR003465; Prot_inh_Pin2.
DR Pfam; PF02428; Prot_inhib_II_2.
KW Signal.
FT SIGNAL 1 31 Potential.
FT SEQUENCE 152 AA; 16432 MW; D86DD64DCD46125C CRC64;
Query Match 71.6%; Score 612; DB 2; Length 152;
Best Local Similarity 74.7%; Pred. No. 1.1e-47;
Matches 109; Conservative 9; Mismatches 24; Indels 4; Gaps 2;
QY 1 MAVHK-VSFLACLVLGWMFLLA---KHVDKAKTREGHFSYGCIPRSEGSPOKPICTN 56
DB 1 MAVHKEVSSLAAYLLVLGLFLFVSAIKHVDKAPCTREGCNLGYGICPRSEGSPEINICTN 60
QY 57 CCGYKGCNYSYSAKGLICEGSDPRNPKDCTFECDTQIAYSKCPRSEGRKMIKPTGCTT 116
DB 61 CCGYKGCNYSYANGTIFCEGSSDPKPNVCPYCDGDIAYSKCPRESGOTIIVPGCTT 120
QY 117 CQTGYQCYTFDQDGFVCEGSPEP 142
DB 121 CQTGYKGCYVFSKGEFVCEGSIEP 146
RESULT 4
AAR37362 PRELIMINARY; PRT; 152 AA.
ID AAR37362;
AC AAR37362;
DT 24-MAY-2004 (TrEMBLrel. 27, Created)
DT 24-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DE Proteinase inhibitor 2b precursor.
GN PIN2B.
OS Solanum nigrum (Black nightshade).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4112;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15078438;
RA Schmidt D.D., Kessler A., Kessler D., Schmidt S., Lim M., Gase K.,
RA Baldwin I.T.;
RT "Solanum nigrum: a model ecological expression system and its tools.";
RL Mol. Ecol. 13:981-995(2004).
DR EMBL; AY422686; AAR37362.1; -.
KW Signal.
FT SIGNAL 1 31 Potential.
FT SEQUENCE 152 AA; 16432 MW; D86DD64DCD46125C CRC64;
Query Match 71.6%; Score 612; DB 2; Length 152;
Best Local Similarity 74.7%; Pred. No. 1.1e-47;
Matches 109; Conservative 9; Mismatches 24; Indels 4; Gaps 2;
QY 1 MAVHK-VSFLACLVLGWMFLLA---KHVDKAKTREGHFSYGCIPRSEGSPOKPICTN 56
DB 1 MAVHKEVSSLAAYLLVLGLFLFVSAIKHVDKAPCTREGCNLGYGICPRSEGSPEINICTN 60
QY 57 CCGYKGCNYSYSAKGLICEGSDPRNPKDCTFECDTQIAYSKCPRSEGRKMIKPTGCTT 116
DB 61 CCGYKGCNYSYANGTIFCEGSSDPKPNVCPYCDGDIAYSKCPRESGOTIIVPGCTT 120
QY 117 CQTGYQCYTFDQDGFVCEGSPEP 142
DB 121 CQTGYKGCYVFSKGEFVCEGSIEP 146
```



```

Db 1 MAVHKEVSSLAYLLVLGLFLPVSIAKHVDKPTRECNGLYGICPRSEGSPEINICTN 60
Qy 57 CCGYKGCNYYSAKGLICSGSDPNPKDCTFECDTQIAYSCKPRSEGMIKPTGCTT 116
Db 61 CCGYKGCNYYSAKGLICSGSDPNPKDCTFECDTQIAYSCKPRSEGMIKPTGCTT 120
Qy 117 CCGYKGCNYYSAKGLICSGSDPNPKDCTFECDTQIAYSCKPRSEGMIKPTGCTT 142
Db 121 CCGYKGCNYYSAKGLICSGSDPNPKDCTFECDTQIAYSCKPRSEGMIKPTGCTT 146

RESULT 5
IP23_LYCES
ID IP23_LYCES STANDARD; PRT; 201 AA.
AC Q43502;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Proteinase inhibitor type II CEVI57 precursor.
GN Name=CEVI57;
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Rutgers; TISSUE=Leaf;
RX MEDLINE=96252900; PubMed=8672818;
RA Gadea J., Mayda E., Conejero V., Vera P.;
RT "Characterization of defense-related genes ectopically expressed in
RT viroid-infected tomato plants.";
RL Mol. Plant Microbe Interact. 9:409-415 (1996).
CC -I- INDUCTION: By viroid infection.
CC -I- SIMILARITY: Belongs to the potato type II proteinase inhibitor
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X94946; CAA64416.1; -.
DR PIR; T07011; T07011.
DR HSSP; P05119; IOYV.
DR InterPro; IPR003465; Prot_inh_Pin2.
DR Pfam; PF02428; Prot_inh_II_3.
KW Repeat; Serine protease inhibitor; Signal.
FT SIGNAL 1 23 Potential.
FT CHAIN 24 201 Proteinase inhibitor type II CEVI57.
FT REPEAT 27 83 1.
FT REPEAT 84 143 2.
FT REPEAT 144 199 3.
FT SITE 32 33 Reactive bond for trypsin (Probable).
FT SITE 147 148 Reactive bond for trypsin (Probable).
FT DISULFID 30 118 By similarity.
FT DISULFID 34 114 By similarity.
FT DISULFID 42 124 By similarity.
FT DISULFID 54 91 By similarity.
FT DISULFID 57 75 By similarity.
FT DISULFID 58 87 By similarity.
FT DISULFID 64 100 By similarity.
FT DISULFID 117 135 By similarity.
SQ SEQUENCE 201 AA; 21419 MW; A3FCAB93773D8590 CRC64;

Query Match 70.4%; Score 602; DB 1; Length 201;
Best Local Similarity 75.7%; Pred. No. 1.le-46;
Matches 109; Conservative 9; Mismatches 24; Indels 2; Gaps 2;

Qy 1 MAVHKEVSSLAYLLVLGLFLPVSIAKHVDKPTRECNGLYGICPRSEGSPEINICTN 59
Qy 60 GYKGCNYYSAKGLICSGSDPNPKDCTFECDTQIAYSCKPRSEGMIKPTGCTT 119
Db 60 GYKGCNYYSAKGLICSGSDPNPKDCTFECDTQIAYSCKPRSEGMIKPTGCTT 119
Qy 120 GYKGCNYYSAKGLICSGSDPNPKDCTFECDTQIAYSCKPRSEGMIKPTGCTT 143
Db 120 GYKGCNYYSAKGLICSGSDPNPKDCTFECDTQIAYSCKPRSEGMIKPTGCTT 146

RESULT 6
IP21_LYCES
ID IP21_LYCES STANDARD; PRT; 148 AA.
AC P05119;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Wound-induced proteinase inhibitor II precursor.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=85207658; PubMed=3838986;
RA Graham J.S., Pearce G., Merryweather J., Titani K., Ericsson L.H.,
RA Ryan C.A.;
RT "Wound-induced proteinase inhibitors from tomato leaves. II. The cDNA-
RT deduced primary structure of pre-inhibitor II.";
RL J. Biol. Chem. 260:6561-6564 (1985).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 31-141 IN COMPLEX WITH
RP SUBTILISIN.
RX PubMed=12684499; DOI=10.1074/jbc.M3020202000;
RA Barrette-Ng I.H., Ng K.K.-S., Cherney M.M., Pearce G., Ryan C.A.,
RA James M.N.G.;
RT "Structural basis of inhibition revealed by a 1:2 complex of the two-
RT headed tomato inhibitor-II and subtilisin Carlsberg.";
RL J. Biol. Chem. 278:24062-24071 (2003).
CC -I- FUNCTION: Potent inhibitor of both trypsin and chymotrypsin.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- INDUCTION: Mechanical damage (i.e., insect chewing) to this plant
CC results in the systemic release of a factor from the wound site.
CC Within the leaves it induces the cytoplasmic synthesis of
CC proteinase inhibitors I and II.
CC -I- SIMILARITY: Belongs to the potato type II proteinase inhibitor
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K03291; AAA34201.1; -.
DR PIR; B24048; B24048.
DR PDB; IOYV; X-ray; I=26-148.
DR PDB; 1PJU; X-ray; A/B/D=26-148.
DR InterPro; IPR003465; Prot_inh_Pin2.
DR Pfam; PF02428; Prot_inh_II_2.
DR 3D-structure; Repeat; Serine protease inhibitor; Signal.
FT SIGNAL 1 25
FT CHAIN 26 148 Wound-induced proteinase inhibitor II.
FT REPEAT 26 81 1 (trypsin-inhibitory).
FT REPEAT 83 141 2 (chymotrypsin-inhibitory).
FT SITE 30 31 Reactive bond for trypsin.
FT SITE 87 88 Reactive bond for chymotrypsin.

```

```
FT DISULFID 28 116
FT DISULFID 32 112
FT DISULFID 40 122
FT DISULFID 52 89
FT DISULFID 55 73
FT DISULFID 56 85
FT DISULFID 62 98
FT DISULFID 115 133
SQ SEQUENCE 148 AA; 16293 MW; 468A2F653971AAFC CRC64;

Query Match 59.7%; Score 596; DB 1; Length 148;
Best Local Similarity 70.3%; Pred. No. 3e-46;
Matches 104; Conservative 15; Mismatches 25; Indels 4; Gaps 2;

QY 1 MAVHK-VSFLACLIVLGWMLLAKHVDKACTRECGHFSYGICPRSEGSPOKPICTNCCS 59
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 1 MAVHKEVNFVAYLLIVLGMLF---YVDKACTRECGNLGFGICPRSEGSPLNINCCS 57
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 60 GYKGCNYSAGDLICEGSDPRNPKDCTFECDTQIAYSKPRSEGMIIKPTGCTTCCT 119
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 58 GYKGCNYSNFGKFCIGESDPRNACTFNCDPNIAYSRCPRSQGSLSIYPTGCTTCCT 117
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 120 GYGCYYFDODGDFVCEGESPEPKTTAY 147
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 118 GYKGCYYFGDKGKFCVCEGESDEPKANNY 145
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 7
AAG12170 PRELIMINARY; PRT; 148 AA.
ID AAG12170;
AC AAG12170;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Proteinase inhibitor II.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae I;
OC Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RA Xie X., Wu N.;
RT "Isolation of tomato proteinase inhibitor II gene and analysis of its
structure."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY007240; AAG12170.1; -.
SQ SEQUENCE 148 AA; 16293 MW; 468A2F653971AAFC CRC64;

Query Match 69.7%; Score 596; DB 2; Length 148;
Best Local Similarity 70.3%; Pred. No. 3e-46;
Matches 104; Conservative 15; Mismatches 25; Indels 4; Gaps 2;

QY 1 MAVHK-VSFLACLIVLGWMLLAKHVDKACTRECGHFSYGICPRSEGSPOKPICTNCCS 59
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 1 MAVHKEVNFVAYLLIVLGMLF---YVDKACTRECGNLGFGICPRSEGSPLNINCCS 57
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 60 GYKGCNYSAGDLICEGSDPRNPKDCTFECDTQIAYSKPRSEGMIIKPTGCTTCCT 119
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 58 GYKGCNYSNFGKFCIGESDPRNACTFNCDPNIAYSRCPRSQGSLSIYPTGCTTCCT 117
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 120 GYGCYYFDODGDFVCEGESPEPKTTAY 147
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 118 GYKGCYYFGDKGKFCVCEGESDEPKANNY 145
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 8
BAC76901 PRELIMINARY; PRT; 148 AA.
ID BAC76901
AC BAC76901;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
SQ SEQUENCE 148 AA; 16293 MW; 468A2F653971AAFC CRC64;

Query Match 69.7%; Score 596; DB 2; Length 148;
Best Local Similarity 70.3%; Pred. No. 3e-46;
Matches 104; Conservative 15; Mismatches 25; Indels 4; Gaps 2;

QY 1 MAVHK-VSFLACLIVLGWMLLAKHVDKACTRECGHFSYGICPRSEGSPOKPICTNCCS 59
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 1 MAVHKEVNFVAYLLIVLGMLF---YVDKACTRECGNLGFGICPRSEGSPLNINCCS 57
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 60 GYKGCNYSAGDLICEGSDPRNPKDCTFECDTQIAYSKPRSEGMIIKPTGCTTCCT 119
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 58 GYKGCNYSNFGKFCIGESDPRNACTFNCDPNIAYSRCPRSQGSLSIYPTGCTTCCT 117
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 120 GYGCYYFDODGDFVCEGESPEPKTTAY 147
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 118 GYKGCYYFGDKGKFCVCEGESDEPKANNY 145
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
```

```
DE Protease inhibitor II.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Moneymaker;
RA Wada M., Matsuda Y., Fujita K., Nanjo A., Kakutani K., Nonomura T.,
RA Matsuda Y., Toyoda H.;
RT "Mature mRNAs in targeted single cells can be amplified by RT-PCR of
nucleus-free cytosolic contents sucked up with micropipette.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB110700; BAC76901.1; -.
KW Protease.
SQ SEQUENCE 148 AA; 16293 MW; 468A2F653971AAFC CRC64;

Query Match 69.7%; Score 596; DB 2; Length 148;
Best Local Similarity 70.3%; Pred. No. 3e-46;
Matches 104; Conservative 15; Mismatches 25; Indels 4; Gaps 2;

QY 1 MAVHK-VSFLACLIVLGWMLLAKHVDKACTRECGHFSYGICPRSEGSPOKPICTNCCS 59
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 1 MAVHKEVNFVAYLLIVLGMLF---YVDKACTRECGNLGFGICPRSEGSPLNINCCS 57
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 60 GYKGCNYSAGDLICEGSDPRNPKDCTFECDTQIAYSKPRSEGMIIKPTGCTTCCT 119
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 58 GYKGCNYSNFGKFCIGESDPRNACTFNCDPNIAYSRCPRSQGSLSIYPTGCTTCCT 117
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 120 GYGCYYFDODGDFVCEGESPEPKTTAY 147
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 118 GYKGCYYFGDKGKFCVCEGESDEPKANNY 145
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 9
AAN05016 PRELIMINARY; PRT; 148 AA.
ID AAN05016
AC AAN05016;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Proteinase inhibitor II.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang H., Xie X., Wu N., Huang M.;
RT "Tomato proteinase inhibitor II genomic DNA including promoter
sequence."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY129402; AAN05016.1; -.
SQ SEQUENCE 148 AA; 16293 MW; 468A2F653971AAFC CRC64;

Query Match 69.7%; Score 596; DB 2; Length 148;
Best Local Similarity 70.3%; Pred. No. 3e-46;
Matches 104; Conservative 15; Mismatches 25; Indels 4; Gaps 2;

QY 1 MAVHK-VSFLACLIVLGWMLLAKHVDKACTRECGHFSYGICPRSEGSPOKPICTNCCS 59
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 1 MAVHKEVNFVAYLLIVLGMLF---YVDKACTRECGNLGFGICPRSEGSPLNINCCS 57
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 60 GYKGCNYSAGDLICEGSDPRNPKDCTFECDTQIAYSKPRSEGMIIKPTGCTTCCT 119
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 58 GYKGCNYSNFGKFCIGESDPRNACTFNCDPNIAYSRCPRSQGSLSIYPTGCTTCCT 117
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

QY 120 GYGCYYFDODGDFVCEGESPEPKTTAY 147
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 118 GYKGCYYFGDKGKFCVCEGESDEPKANNY 145
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
```

```

RESULT 11
Q84L56 PRELIMINARY; PRT; 154 AA.
AC Q84L56;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 05-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE Proteinase inhibitor II protein (Proteinase inhibitor 2 protein).
GN NamePIN-I12x; Synonyms=Pin-I12x;
OS Solanum phureja.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=172790;
ON [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Wound-induced leaf;
RC Bu Q., Yang S., Wan J.;
RA Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RL [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Wound-induced leaf;
RC Bu Q., Wu L., Yang S., Wan J.;
RA Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL: AY247794; AAC8244.1; -.
DR EMBL: AY517498; AAS01731.1; -.
DR HSP; P05119; IOYV.
DR GO; GO:0004867; F:serine-type endopeptidase inhibitor activity; IEA.
DR InterPro; IPR003465; Prot_inh_Pin2.
DR Pfam; PF02428; Prot_inhib_II_2.
DR SQ SEQUENCE 154 AA; 16641 MW; F7019E715BF58553 CRC64;

Query Match 69.6%; Score 595; DB 2; Length 154;
Best Local Similarity 70.9%; Pred. No. 3.8e-46;
Matches 107; Conservative 15; Mismatches 25; Indels 4; Gaps 3;

Qy 1 MAVHK-VSFACLL-VLGMWFLA--KHVDKACTRECQHFSGYICPRSEGSPOKPICTN 56
Db 1 MAVKVENFVAYLLVLGLLVLSAMEHVDKAKATLECGNLGFGICPRSEGSPENRICTN 60
Qy 57 CCSGYKGCNYSAGKDLICEGSDPRPKDCTECDFQIAYSKCPRSEGMIIKPTGCTT 116
Db 61 CCAGYKGCNYSANGAFICEGSDPKPNPCRNCDTNIAYSKCPRSEGSGLIYPTGCTT 120
Qy 117 CCTGYOCYFYDDQDGFVCEGSEPEPKTAY 147
Db 121 CCTGYKCYFYGKNGKPFVCEGSEDEPKANNY 151

RESULT 12
AAS01731 PRELIMINARY; PRT; 154 AA.
ID AAS01731
AC AAS01731;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Proteinase inhibitor 2 protein.
OS Solanum phureja.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=172790;
ON [1]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. IVP101; TISSUE=Wound-induced leaf;
RC Bu Q., Wu L., Yang S., Wan J.;
RA "diploid potato proteinase inhibitor.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL: AY517498; AAS01731.1; -.
DR SQ SEQUENCE 154 AA; 16641 MW; F7019E715BF58553 CRC64;

```

```

Query Match      69.6%; Score 595; DB 2; Length 154;
Best Local Similarity 70.9%; Pred. No. 3.8e-46;
Matches 107; Conservative 15; Mismatches 25; Indels 4; Gaps 3;

QY 1 MAVHK-VSFLACLL-VLGMFLLA--KHVDAKACTRECCHFSYICPRSEGSQKPICTN 56
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MAVHKEVNFVAYLLI-VLGLVLVLSAMEHVDKACTLECGNLGFGICPRSEGSPEINICTN 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 57 CCGYKGCNYYSAKGDILICGESDPRNPDKCTFECDTQIAYSKCPRSEGKMIKPTGCTT 116
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 CCAGYKGCNYYSAKGFICGESDPRNPVCPNCDNIAYSKCPRSEGKSLIYPTGCTT 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 117 CTGYQCYYFDQDGFVCEGESPEPKTAY 147
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 CCTGYKGCYYFGKNGKFCVCEGESDEPKANMY 151
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
IP2X_SOLTU STANDARD; PRT; 158 AA.
AC Q007B2;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Proteinase inhibitor type II precursor.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Russet Burbank;
RA Choi Y., Moon Y., Lee J.S.;
RT "Primary structure of two proteinase inhibitor II genes closely linked
in the potato genome.";
RL Korean J. Biochem. 23:214-220(1990).
CC -1- SIMILARITY: Belongs to the potato type II proteinase inhibitor
family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; Z12753; CAA78277.1; -.
CC DR PIR; T07597; T07597.
CC DR HSSP; P01080; 4SGB.
CC DR InterPro; IPR003465; Prot_inh_Pin2.
CC DR Pfam; PF02428; Prot_inh_II_3.
CC KW Repeat; Serine protease inhibitor; Signal.
CC SIGNAL 1 24 Potential.
CC FT CHAIN 25 158 Proteinase inhibitor type II.
CC FT REPEAT 29 86 1.
CC FT REPEAT 87 146 2.
CC FT SITE 35 36 Reactive bond for chymotrypsin
(Possible).
CC FT SITE 92 93 Reactive bond for trypsin (Probable).
CC FT DISULFID 33 121 By similarity.
CC FT DISULFID 37 117 By similarity.
CC FT DISULFID 45 127 By similarity.
CC FT DISULFID 57 94 By similarity.
CC FT DISULFID 60 78 By similarity.
CC FT DISULFID 61 90 By similarity.
CC FT DISULFID 67 103 By similarity.
CC FT DISULFID 120 138 By similarity.
CC SQ SEQUENCE 158 AA; C84DF44B015F248E CRC64;

Query Match      69.4%; Score 593.5; DB 1; Length 158;
Best Local Similarity 71.9%; Pred. No. 5.3e-46;

QY 1 MAVHK-VSFLACLL-VLGMFLLA--KHVDAKACTRECCHFSYICPRSEGSQKPICTN 56
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MAVHKEVNFVAYLLI-VLGLVLVLSAMEHVDKACTLECGNLGFGICPRSEGSPEINICTN 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 58 CCGYKGCNYYSAKGDILICGESDPRNPDKCTFECDTQIAYSKCPRSEGKMIKPTGCTT 117
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 CAGYKGCNYYSAKGTFCGQSHPKPKACPRNCDPHAIYSKCPRSGGKTLIYPTGCTT 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 118 CTGYQCYYFDQDGFVCEGESPEPK 143
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 CTGYTDCYYFGKDGKFCVCEGESIEPK 146
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 14
IP25_SOLTU STANDARD; PRT; 154 AA.
AC Q41488;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Proteinase inhibitor type II P303.51 precursor.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Bintje; TISSUE=Tuber;
RA Jongma M.A., Bakker P.L., Stiekema W.J., Bosch D.D.;
RT "Phase display of a double-headed proteinase inhibitor: analysis of
the binding domains of potato proteinase inhibitor II.";
RL Mol. Breed. 1:181-191(1995).
CC -1- SIMILARITY: Belongs to the potato type II proteinase inhibitor
family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; L37519; AAS3278.1; -.
CC DR HSSP; P05119; 1OYV.
CC DR InterPro; IPR003465; Prot_inh_Pin2.
CC DR Pfam; PF02428; Prot_inh_II_2.
CC KW Repeat; Serine protease inhibitor; Signal.
CC SIGNAL 1 25 Potential.
CC FT CHAIN 26 154 Proteinase inhibitor type II P303.51.
CC FT REPEAT 31 87 1.
CC FT REPEAT 88 147 2.
CC FT SITE 36 37 Reactive bond for chymotrypsin
(Possible).
CC FT SITE 93 94 Reactive bond for trypsin (Probable).
CC FT DISULFID 34 122 By similarity.
CC FT DISULFID 38 118 By similarity.
CC FT DISULFID 46 128 By similarity.
CC FT DISULFID 58 95 By similarity.
CC FT DISULFID 61 79 By similarity.
CC FT DISULFID 62 91 By similarity.
CC FT DISULFID 68 104 By similarity.
CC FT DISULFID 121 139 By similarity.
CC SQ SEQUENCE 154 AA; 16660 MW; AF0BFD1F26B6F224 CRC64;

Query Match      68.9%; Score 589; DB 1; Length 154;
Best Local Similarity 70.2%; Pred. No. 1.3e-45;
Matches 106; Conservative 15; Mismatches 26; Indels 4; Gaps 3;

QY 1 MAVHK-VSFLACLL-VLGMFLLA--KHVDAKACTRECCHFSYICPRSEGSQKPICTN 56

```

Db	58	GYKGCNYSAFGFCEGSDPKPKACPLNCDNIAYSRCPSEKSLIYPTGCTCT	117
Qy	120	GYQGCYYFDQDQDFVCEGESPEPK	143
Db	118	GYKGCYYFTGNGKFCVCEGESDEPK	141

Search completed: January 24, 2005, 16:52:05
Job time : 194 secs

RESULT 15

IP2Y_SOLTU	IP2Y_SOLTU	STANDARD;	PRT;	147 AA.
ID	IP2Y_SOLTU	STANDARD;	PRT;	147 AA.
AC	Q41489;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Proteinase inhibitor type II precursor.			
DE	Solanum tuberosum (Potato).			
OS	Solanum tuberosum (Potato).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;			
OC	lamids; Solanales; Solanaceae; Solanum.			
OC	NCBI_TaxID=4113;			
RN	[1]			
RP	SEQUENCE FROM N. A.			
RC	STRAIN=cv, Russet Burbank;			
RC	Choi Y., Kim J.W., Lee J.S.;			
RL	Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.			
CC	-I- SIMILARITY: Belongs to the potato type II proteinase inhibitor			
CC	family.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isg-sib.ch/announce/>) or send an email to license@sib-sib.ch.

CC	EMBL; Z11992; CAA78383.1; --	
DR	PIR; S24973; S24973.	
DR	HSSP; P05119; IOYV.	
DR	InterPro; IPR003465; Prot_inh_Pin2.	
DR	PFam; PF02428; Prot_inhib_II; 2.	
KW	Repeat; Serine protease inhibitor; Signal.	
KW	Repeat; Serine protease inhibitor; Signal.	
FT	SIGNAL	1 25
FT	CHAIN	26 147
FT	REPEAT	25 81
FT	REPEAT	82 141
FT	REPEAT	80 31
FT	SITE	87 88
FT	SITE	87 88
FT	DISULFID	28 116
FT	DISULFID	32 112
FT	DISULFID	40 122
FT	DISULFID	52 89
FT	DISULFID	55 73
FT	DISULFID	56 85
FT	DISULFID	62 98
FT	DISULFID	115 133
FT	DISULFID	115 133
SO	SEQUENCE	147 AA: 15936 MW: 703456551B54F968 CRC64:

Query Match	68.5%	Score	586;	DB 1;	Length	147;	
Best Local Similarity	72.2%	Pred. No.	2.4e-45;				
Matches 104: Conservative	12;	Mismatches	24;	Indels	4;	Gaps	2;

1	MAVHK-VSPFLACLLVLGMFWFLAKHVDVDAKACTRCGHFSYGI	CPRSEGSQPKICTNCCS	59
	: : : : : : :		
	1	MAVHKVSVFVLLIIVLGMFL---	YVDALGCTRCGNLGFGI
			CPRSEGSTPNPICNCS
			57
60	GYKGCNYSYSAKGDLICEGESDPRNPVKCTFCDTDQIAYSKPRSEGMWIKPTGCTTCCT		119

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2005, 16:44:21 ; Search time 40 Seconds
(without alignments)
245.377 Million cell updates/sec

Title: US-10-725-829-2

Perfect score: 855

Sequence: 1 MAVHKVSFLACLLVLGWMFL.....QDGFVCEGSEPEPKTTAYF 148

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/protdata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/protdata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/protdata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/protdata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/protdata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/protdata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	379	44.3	368	3	US-08-454-295-3
2	379	44.3	368	3	US-09-431-500A-3
3	379	44.3	368	4	US-09-431-498-3
4	379	44.3	368	4	US-09-431-499-3
5	224	26.2	58	3	US-08-454-295-5
6	224	26.2	58	3	US-09-431-500A-5
7	224	26.2	58	4	US-09-431-498-5
8	224	26.2	58	4	US-09-431-499-5
9	204	23.9	58	3	US-08-454-295-6
10	204	23.9	58	3	US-09-431-500A-6
11	204	23.9	58	4	US-09-431-498-6
12	204	23.9	58	4	US-09-431-499-6
13	203	23.7	58	3	US-08-454-295-7
14	203	23.7	58	3	US-08-454-295-8
15	203	23.7	58	3	US-09-431-500A-7
16	203	23.7	58	3	US-09-431-500A-8
17	203	23.7	58	4	US-09-431-498-7
18	203	23.7	58	4	US-09-431-498-8
19	203	23.7	58	4	US-09-431-499-7
20	203	23.7	58	4	US-09-431-499-8
21	199	23.3	58	3	US-08-454-295-9
22	199	23.3	58	3	US-09-431-500A-9
23	199	23.3	58	4	US-09-431-498-9
24	199	23.3	58	4	US-09-431-499-9
25	120	14.0	54	3	US-08-454-295-10
26	120	14.0	54	3	US-09-431-500A-10
27	120	14.0	54	4	US-09-431-498-10

28 120 14.0 54 4 US-09-431-499-10 Sequence 10, Appli
29 95.5 11.2 578 3 US-08-981-392-13 Sequence 13, Appli
30 95.5 11.2 578 4 US-09-908-322-13 Sequence 13, Appli
31 94 11.0 148 3 US-08-882-907-15 Sequence 15, Appli
32 94 11.0 148 4 US-10-032-658-15 Sequence 15, Appli
33 92.5 10.8 2321 4 US-09-230-652-2 Sequence 2, Appli
34 91.5 10.7 242 4 US-09-312-283C-393 Sequence 193, App
35 91.5 10.7 420 4 US-09-907-794A-109 Sequence 109, App
36 91.5 10.7 420 4 US-09-905-125A-109 Sequence 109, App
37 91.5 10.7 420 4 US-09-902-775A-137 Sequence 109, App
38 91.5 10.7 420 4 US-09-906-700-109 Sequence 290, App
39 91.5 10.7 420 4 US-10-140-002-290 Sequence 109, App
40 91.5 10.7 420 4 US-09-903-603A-109 Sequence 109, App
41 91 10.6 1417 3 US-08-900-230-3 Sequence 3, Appli
42 89.5 10.5 314 2 US-08-460-309-19 Sequence 19, Appli
43 89.5 10.5 314 2 US-08-125-077-19 Sequence 19, Appli
44 89.5 10.5 2213 1 US-08-727-034-3 Sequence 3, Appli
45 89.5 10.5 2214 1 US-08-727-034-7 Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-08-454-295-3
; Sequence 3, Application US/08454295
; Patent No. 6031087

GENERAL INFORMATION:

APPLICANT: Anderson, Marilyn A.
APPLICANT: Atkinson, Angela H.
APPLICANT: Heath, Robyn L.
APPLICANT: Clarke, Adrienne E.
TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States of America
ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,295
FILING DATE: 01-SEP-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9748
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 368 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-454-295-3

Query Match 44.3%; Score 379; DB 3; Length 368;

Best Local Similarity 58.1%; Pred. No. 2.3e-27;

Matches 68; Conservative 9; Mismatches 36; Indels 4; Gaps 3;

Qy 28 KACTRECG-HFSGICPRSEGSQKPCICTNCCSGYKCNYSKAGDLICEGSDPRPKD 86

Db 1 KACTLNCDPRIAYGVCPREKKNDRICNCCAGTKGCKYFSDDTFVCEGSDPRPKA 60


```
; MOLECULE TYPE: protein
US-09-431-499-3

Query Match      44.3%; Score 379; DB 4; Length 368;
Best Local Similarity 58.1%; Pred. No. 2.3e-27;
Matches 68; Conservative 9; Mismatches 36; Indels 4; Gaps 3;

Qy 28 KACTRCG-HFSGICPRSGSQPKICTNCCSGYKCNYSKAGDLICGEGSDPRNPKD 86
Db 1 KACTLNCDPRIAYGVCPRSEKKNDRICTNCCAGTKGCKYFSDGTFVCEGESDPRNPKA 60

Qy 87 CTPECCTQIAYSKPRSEGMIIKPTGCTTCTCGYGCYCFDQDGFVCEGESPEPK 143
Db 61 CTLNCDPRIAYGVCPRSEKK--NDRICTNCCAGTKGCKYFSDGTFVCEGES-DPR 114

RESULT 5
US-08-454-295-5
; Sequence 5, Application US/08454295
; Patent No. 6031087
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-454-295-5

Query Match      26.2%; Score 224; DB 3; Length 58;
Best Local Similarity 68.5%; Pred. No. 6.9e-14;
Matches 37; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

Qy 53 ICTNCCSGYKCNYSKAGDLICGEGSDPRNPKDCTFECDTQIAYSKCPRSECK 106
Db 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPKACTLNCDPRIAYGVCPRSEK 56

RESULT 6
US-09-431-500A-5
; Sequence 5, Application US/09431500A
; Patent No. 6261821
```

```
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,498
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/454,295
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

Query Match      26.2%; Score 224; DB 3; Length 58;
Best Local Similarity 68.5%; Pred. No. 6.9e-14;
Matches 37; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

Qy 53 ICTNCCSGYKCNYSKAGDLICGEGSDPRNPKDCTFECDTQIAYSKCPRSECK 106
Db 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPKACTLNCDPRIAYGVCPRSEK 56

RESULT 7
US-09-431-498-5
; Sequence 5, Application US/09431498
; Patent No. 6440727
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,498
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/454,295
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

US-09-431-500A-5
Query Match      26.2%; Score 224; DB 3; Length 58;
Best Local Similarity 68.5%; Pred. No. 6.9e-14;
Matches 37; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

Qy 53 ICTNCCSGYKCNYSKAGDLICGEGSDPRNPKDCTFECDTQIAYSKCPRSECK 106
Db 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPKACTLNCDPRIAYGVCPRSEK 56
```

```
; MOLECULE TYPE: protein
US-09-431-498-5

Query Match      26.2%; Score 224; DB 4; Length 58;
Best Local Similarity 68.5%; Pred. No. 6.9e-14;
Matches 37; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGDLICEGSDPRNPDKCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPACTLNCDPRIAYGVCPRSEK 56

RESULT 8
US-09-431-499-5
; Sequence 5, Application US/09431499
; Patent No. 6451573
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF
; TITLE OF INVENTION: AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 01-SEP-1995
; APPLICATION NUMBER: US/09/431,499
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4366
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-431-499-5

Query Match      26.2%; Score 224; DB 4; Length 58;
Best Local Similarity 68.5%; Pred. No. 6.9e-14;
Matches 37; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGDLICEGSDPRNPDKCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPACTLNCDPRIAYGVCPRSEK 56

RESULT 9
US-08-454-295-6
; Sequence 6, Application US/08454295
; Patent No. 6031087
```

```
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 01-SEP-1995
; APPLICATION NUMBER: US/08/454,295
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-454-295-6

Query Match      23.9%; Score 204; DB 3; Length 58;
Best Local Similarity 63.0%; Pred. No. 4.8e-12;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGDLICEGSDPRNPDKCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPACTLNCDPRIAYGVCPRSEK 56

RESULT 10
US-09-431-500A-6
; Sequence 6, Application US/09431500A
; Patent No. 6261821
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; FILE REFERENCE: 9748B
; CURRENT APPLICATION NUMBER: US/09/431,500A
; CURRENT FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Nicotiana glauca
US-09-431-500A-6
```

```
Query Match      23.9%; Score 204; DB 3; Length 58;
Best Local Similarity 63.0%; Pred. No. 4.8e-12;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGKDLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 ICTNCCAGTKGCKYFSDGDTFVCEGSDPRNPKACPRNCDPRIAYGICPLAEK 56

RESULT 11
US-09-431-498-6
; Sequence 6, Application US/09431498
; Patent No. 6440727
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,498
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/454,295
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-431-498-6

Query Match      23.9%; Score 204; DB 4; Length 58;
Best Local Similarity 63.0%; Pred. No. 4.8e-12;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGKDLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 ICTNCCAGTKGCKYFSDGDTFVCEGSDPRNPKACPRNCDPRIAYGICPLAEK 56

RESULT 12
US-09-431-499-6
; Sequence 6, Application US/09431499
; Patent No. 6451573
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.

Query Match      23.9%; Score 204; DB 4; Length 58;
Best Local Similarity 63.0%; Pred. No. 4.8e-12;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGKDLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 ICTNCCAGTKGCKYFSDGDTFVCEGSDPRNPKACPRNCDPRIAYGICPLAEK 56

RESULT 13
US-08-454-295-7
; Sequence 7, Application US/08454295
; Patent No. 6031087
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,499
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-431-499-6

Query Match      23.9%; Score 204; DB 4; Length 58;
Best Local Similarity 63.0%; Pred. No. 4.8e-12;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGKDLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 ICTNCCAGTKGCKYFSDGDTFVCEGSDPRNPKACPRNCDPRIAYGICPLAEK 56
```

```
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/431,499
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-431-499-6

Query Match      23.9%; Score 204; DB 4; Length 58;
Best Local Similarity 63.0%; Pred. No. 4.8e-12;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGKDLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 ICTNCCAGTKGCKYFSDGDTFVCEGSDPRNPKACPRNCDPRIAYGICPLAEK 56

RESULT 13
US-08-454-295-7
; Sequence 7, Application US/08454295
; Patent No. 6031087
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

US-08-454-295-8

Search completed: January 24, 2005, 16:53:33
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2005, 16:47:41 ; Search time 55 Seconds
(without alignments)
972.197 Million cell updates/sec

Title: US-10-725-829-2
Perfect score: 855
Sequence: 1 MAVHKVSFLACLLVLGMFL.....QGDGFCVCEGSEPEPKTTAYF 148

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCT05_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	855	100.0	148	17	US-10-725-829-2
2	615	71.9	152	17	US-10-725-829-4
3	379	44.3	368	10	US-09-812-502-3
4	379	44.3	368	14	US-10-164-961-3
5	379	44.3	368	14	US-10-157-622-3
6	379	44.3	368	14	US-10-072-809A-57
7	224	26.2	58	10	US-09-812-502-5
8	224	26.2	58	14	US-10-164-961-5
9	224	26.2	58	14	US-10-157-622-5
10	204	23.9	58	10	US-09-812-502-6
11	204	23.9	58	14	US-10-164-961-6
12	204	23.9	58	14	US-10-157-622-6
13	203	23.7	58	10	US-09-812-502-7

14	203	23.7	58	10	US-09-812-502-8	Sequence 8, Appli
15	203	23.7	58	14	US-10-164-961-7	Sequence 7, Appli
16	203	23.7	58	14	US-10-164-961-8	Sequence 8, Appli
17	203	23.7	58	14	US-10-157-622-7	Sequence 7, Appli
18	203	23.7	58	14	US-10-157-622-8	Sequence 8, Appli
19	199	23.3	58	10	US-09-812-502-9	Sequence 9, Appli
20	199	23.3	58	14	US-10-164-961-9	Sequence 9, Appli
21	199	23.3	58	14	US-10-157-622-9	Sequence 9, Appli
22	127	14.9	50	10	US-09-991-209-13	Sequence 13, Appli
23	120	14.0	54	10	US-09-812-502-10	Sequence 10, Appli
24	120	14.0	54	14	US-10-164-961-10	Sequence 10, Appli
25	120	14.0	54	14	US-10-157-622-10	Sequence 10, Appli
26	101.5	11.9	78	17	US-10-425-115-349792	Sequence 349792
27	100.5	11.8	1057	14	US-10-189-971-6	Sequence 6, Appli
28	100.5	11.8	1192	14	US-10-189-971-18	Sequence 18, Appli
29	100.5	11.8	1207	14	US-10-189-971-20	Sequence 20, Appli
30	100.5	11.8	1251	14	US-10-189-971-16	Sequence 16, Appli
31	100.5	11.8	1342	14	US-10-189-971-24	Sequence 24, Appli
32	100.5	11.8	1477	14	US-10-189-971-8	Sequence 8, Appli
33	100.5	11.8	1512	14	US-10-189-971-10	Sequence 10, Appli
34	100.5	11.8	1535	14	US-10-189-971-14	Sequence 14, Appli
35	100.5	11.8	1570	14	US-10-189-971-12	Sequence 12, Appli
36	100.5	11.8	1593	14	US-10-189-971-4	Sequence 4, Appli
37	100.5	11.8	1617	14	US-10-184-644-235	Sequence 235, App
38	100.5	11.8	1617	14	US-10-184-634-235	Sequence 235, App
39	100.5	11.8	1617	14	US-10-063-685-61	Sequence 61, Appli
40	100.5	11.8	1628	14	US-10-189-971-2	Sequence 2, Appli
41	99	11.6	214	16	US-10-767-701-36663	Sequence 36663, A
42	98	11.5	79	17	US-10-425-115-349791	Sequence 349791,
43	96.5	11.3	1574	10	US-09-825-751A-77	Sequence 77, Appli
44	95.5	11.2	578	9	US-09-908-322-13	Sequence 13, Appli
45	95.5	11.2	578	10	US-09-783-931-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1
US-10-725-829-2
; Sequence 2, Application US/10725829
; Publication No. US20040205846A1
; GENERAL INFORMATION:
; APPLICANT: Chye, Mee Len
; APPLICANT: Xu, Zeng-Fu
; APPLICANT: Sin, Suk Fong
; TITLE OF INVENTION: Genetically Modified Plants with Heterogenous Proteinase Inhibitor
; TITLE OF INVENTION: SAPI2A or SAPI2B and Methods of Use Thereof for the Inhibition
; TITLE OF INVENTION: of Trypsin- and Chymotrypsin-Like Activities
; FILE REFERENCE: 9661-043-999
; CURRENT APPLICATION NUMBER: US/10/725,829
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: 60/429,992
; PRIOR FILING DATE: 2002-11-29
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Solanum americanum
US-10-725-829-2

Query Match	100.0%;	Score 855;	DB 17;	Length 148;
Best Local Similarity	100.0%;	Pred. No. 1.8e-71;		
Matches 148;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	MAVHKVSFLACLLVLGMFLAKHVDKACTRECGHFSYGCIPRSGSPQKPICTNCCSG	60	
Db	1	MAVHKVSFLACLLVLGMFLAKHVDKACTRECGHFSYGCIPRSGSPQKPICTNCCSG	60	
Qy	61	YKGCNYSAGDLICBESDPRNPDKCTPECDTQIAYSKPRSEGKMIKPTGCTTCCTG	120	
Db	61	YKGCNYSAGDLICBESDPRNPDKCTPECDTQIAYSKPRSEGKMIKPTGCTTCCTG	120	

QY 121 YQCYFYFDQDGFVCEGESPEPKTAYF 148
Db 121 YQCYFYFDQDGFVCEGESPEPKTAYF 148
RESULT 2
US-10-725-829-4
; Sequence 4, Application US/10725829
; Publication No. US20040205846A1
; GENERAL INFORMATION:
; APPLICANT: Chye, Mee Len
; APPLICANT: Xu, Zeng-Fu
; APPLICANT: Sin, Suk Fong
; TITLE OF INVENTION: Genetically Modified Plants with Heterogenous Proteinase Inhibitor
; TITLE OF INVENTION: SAPIN2A or SAPIN2B and Methods of Use Thereof for the Inhibition
; TITLE OF INVENTION: of Trypsin- and Chymotrypsin-like Activities
; FILE REFERENCE: 9661-043-999
; CURRENT APPLICATION NUMBER: US/10/725,829
; CURRENT FILING DATE: 2003-12-01
; PRIOR FILING DATE: 2002-11-29
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Solanum americanum
US-10-725-829-4

Query Match 71.9%; Score 615; DB 17; Length 152;
Best Local Similarity 75.3%; Pred. No. 3.3e-49;
Matches 110; Conservative 8; Mismatches 24; Indels 4; Gaps 2;
QY 1 MAVHK-VSFLACLVLGMFL---LAKGVDAKACTRECGHFSYIGICPRSGSPQKPICTN 56
Db 1 MAVHKEVSSLAYLLVLGLMFLHVSVAVKHVDKPTRECGNLGYGICPRSGSPENPICTN 60
QY 57 CCSGYGKCNVYSAGKDLICGESDPRNPDKCTPECDTQIAYSKCPRSEGKMIKPTGCTT 116
Db 61 CCSGYGKCNVYSANGYFICGSSDPKNPNTCLFCDGDIAYSKCPRSEGETIIVPTGCTT 120
QY 117 CTGYGQCYFYFDQDGFVCEGESPEP 142
Db 121 CTGYGQCYFYFSEGEFVCEGESDEP 146

RESULT 3
US-09-812-502-3
; Sequence 3, Application US/09812502
; Publication No. US20030027303A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; TITLE OF INVENTION: SEQUENCES ENCODING SAME
; FILE REFERENCE: 9748B
; CURRENT APPLICATION NUMBER: US/09/812,502
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US/09/431,500
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Nicotiana glauca
US-09-812-502-3

Query Match 44.3%; Score 379; DB 10; Length 368;

Best Local Similarity 58.1%; Pred. No. 5.8e-27;
Matches 68; Conservative 9; Mismatches 36; Indels 4; Gaps 3;
QY 28 KACTRECG-HFSYGICPRSGSPQKPICTNCCSGYKCNVYSAGKDLICGESDPRNPDK 86
Db 1 KACTLNCDPRIAYGVCPRSEKKNDRICNCCAGTGGCKYFSDGTFVCEGESDPRNPKA 60
QY 87 CTPECDTQIAYSKCPRSEGKMIKPTGCTTCTGYGQCYFYFDQDGFVCEGESPEP 143
Db 61 CTLNCDPRIAYGVCPRSEKKNDRICNCCAGTGGCKYFSDGTFVCEGES-DPR 114

RESULT 4
US-10-164-961-3
; Sequence 3, Application US/10164961
; Publication No. US20030096388A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/10/164,961
; FILING DATE: 07-Jun-2002
; CLASSIFICATION: <Unknown>
; APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 368 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-164-961-3

Query Match 44.3%; Score 379; DB 14; Length 368;
Best Local Similarity 58.1%; Pred. No. 5.8e-27;
Matches 68; Conservative 9; Mismatches 36; Indels 4; Gaps 3;
QY 28 KACTRECG-HFSYGICPRSGSPQKPICTNCCSGYKCNVYSAGKDLICGESDPRNPDK 86
Db 1 KACTLNCDPRIAYGVCPRSEKKNDRICNCCAGTGGCKYFSDGTFVCEGESDPRNPKA 60
QY 87 CTPECDTQIAYSKCPRSEGKMIKPTGCTTCTGYGQCYFYFDQDGFVCEGESPEP 143
Db 61 CTLNCDPRIAYGVCPRSEKKNDRICNCCAGTGGCKYFSDGTFVCEGES-DPR 114

RESULT 5

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 57
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Nicotiana glauca
US-10-072-809A-57

Query Match      44.3%; Score 379; DB 14; Length 368;
Best Local Similarity 58.1%; Pred. No. 5.8e-27;
Matches 68; Conservative 9; Mismatches 36; Indels 4; Gaps 3;

QY 28 KACTREG-HFSGVIGIPRSGSGPQKPICTNCCSGYKGCNYSYSAKGLDICEGSDPRNPKD 86
    ||||| :|||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 KACTLNCDPRIAYGVCPRSEKKNDRICNCCAGTKGCKYFSDDTFVCEGSDPRNPKA 60
    ||||| :|||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 87 CTPECDTQIAYSKCPRSEGHMIIKPTGCTTCTGYQGYCYFYDQDGFVCEGSGPEPK 143
    ||||| :|||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 CTLNCDPRIAYGVCPRSEKK--NDRICNCCAGTKGCKYFSDDTFVCEGSG-DPR 114
    ||||| :|||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-09-812-502-5
; Sequence 5, Application US/09812502
; Publication No. US20030027303A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; TITLE OF INVENTION: SEQUENCES ENCODING SAME
; FILE REFERENCE: 9748B
; CURRENT APPLICATION NUMBER: US/09/812,502
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US/09/431,500
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Nicotiana glauca
US-09-812-502-5

Query Match      26.2%; Score 224; DB 10; Length 58;
Best Local Similarity 68.3%; Pred. No. 2.1e-13;
Matches 37; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGCNYSYSAKGLDICEGSDPRNPKDCTPECDTQIAYSKCPRSEK 106
    ||||| :|||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3 ICTNCCAGTKGCKYFSDDTFVCEGSDPRNPKACTLNCDPRIAYGVCPRSEK 56
    ||||| :|||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-10-164-961-5
; Sequence 5, Application US/10164961
; Publication No. US20030096388A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530

```



```
;
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/164,961
; FILING DATE: 07-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; MOLECULE TYPE: protein
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
; US-10-164-961-6
;
; Query Match 23.9%; Score 204; DB 14; Length 58;
; Best Local Similarity 63.0%; Pred. No. 1.5e-11;
; Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;
;
; QY 53 ICTNCCSGYKGCNYYSAGKDLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
; |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
; Db 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPKACPRNCDPRIAYGICPLAEK 56
;
; RESULT 12
; US-10-157-622-6
; Sequence 6, Application US/10157622
; Publication No. US20030129720A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; ATKINSON, ANGELA H.
; HEATH, ROBYN L.
; CLARKE, ADRIENNE E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/157,622
; FILING DATE: 29-May-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; MOLECULE TYPE: protein
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
; US-10-164-961-6
;
; Query Match 23.9%; Score 204; DB 14; Length 58;
; Best Local Similarity 63.0%; Pred. No. 1.5e-11;
; Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;
;
; QY 53 ICTNCCSGYKGCNYYSAGKDLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
; |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
; Db 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPKACPRNCDPRIAYGICPLAEK 56
;
; RESULT 13
; US-09-812-502-7
; Sequence 7, Application US/09812502
; Publication No. US20030027303A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; TITLE OF INVENTION: SEQUENCES ENCODING SAME
; FILE REFERENCE: 9748B
; CURRENT APPLICATION NUMBER: US/09/812,502
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US/09/431,500
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Nicotiana glauca
; US-09-812-502-7
;
; Query Match 23.7%; Score 203; DB 10; Length 58;
; Best Local Similarity 63.0%; Pred. No. 1.9e-11;
; Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;
;
; QY 53 ICTNCCSGYKGCNYYSAGKDLICEGSDPRNPKDCTFECDTQIAYSKCPRSEK 106
; |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
; Db 3 ICTNCCAGTKGCKYFSDGTFVCEGESDPRNPKACPRNCDPRIAYGICPLSEK 56
;
; RESULT 14
; US-09-812-502-8
; Sequence 8, Application US/09812502
; Publication No. US20030027303A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; APPLICANT: Atkinson, Angela H.
; APPLICANT: Heath, Robyn L.
; APPLICANT: Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; TITLE OF INVENTION: SEQUENCES ENCODING SAME
; FILE REFERENCE: 9748B
```

; CURRENT APPLICATION NUMBER: US/09/812,502
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US/09/431,500
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 08/454,295
; PRIOR FILING DATE: 1995-09-01
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Nicotiana glauca
US-09-812-502-8

Query Match 23.7%; Score 203; DB 10; Length 58;
Best Local Similarity 63.0%; Pred. No. 1.9e-11;
Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 53 ICTNCCGYKGCNYYSAKGLICEGSDPRNPKDCTFECDTQIAYSKCPRSECK 106
|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 3 ICTNCCAGKKGCKYFSDDDGTFVCEGESDPKNPKACPRNCDGRIAYGICPLSECK 56

RESULT 15
US-10-164-961-7
; Sequence 7, Application US/10164961
; Publication No. US20030096388A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Marilyn A.
; Atkinson, Angela H.
; Heath, Robyn L.
; Clarke, Adrienne E.
; TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/164,961
; FILING DATE: 07-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,295
; FILING DATE: 01-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9748
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-164-961-7

Query Match 23.7%; Score 203; DB 14; Length 58;
Best Local Similarity 63.0%; Pred. No. 1.9e-11;

Matches 34; Conservative 5; Mismatches 15; Indels 0; Gaps 0;
QY 53 ICTNCCGYKGCNYYSAKGLICEGSDPRNPKDCTFECDTQIAYSKCPRSECK 106
|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 3 ICTNCCAGKKGCKYFSDDDGTFVCEGESDPKNPKACPRNCDGRIAYGICPLSECK 56

Search completed: January 24, 2005, 16:54:34
Job time : 56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 24, 2005, 16:41:10 ; Search time 67 Seconds
(without alignments)
792.417 Million cell updates/sec

Title: US-10-725-829-2
Perfect score: 855
Sequence: 1 MAVHKVSFLACLLVLGVML.....QDGFVCGESPBPCKTAYF 148

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	855	100.0	148	ADP02855	Solanum a
2	615	71.9	152	ADP02857	Solanum a
3	585.5	68.5	153	AAR20099	Wound-ind
4	379	44.3	368	AAR54135	Deducted s
5	379	44.3	368	ADI56864	Plant def
6	224	26.2	58	AAR54125	Sequence
7	206	24.1	58	AAR54129	Sequence
8	204	23.9	58	AAR54126	Sequence
9	203	23.7	58	AAR54127	Sequence
10	203	23.7	58	AAR54128	Sequence
11	127	14.9	50	ADG69883	Potato pr
12	118	13.8	54	AAR54130	Sequence
13	116.5	13.6	49	ADN88896	Potato pr
14	106	12.4	444	ADP31478	Human sec
15	104.5	12.2	186	ADP30830	Human sec
16	104	12.2	137	ABM74314	DNA clone
17	100.5	11.8	1057	6 ABP97370	Human kie
18	100.5	11.8	1192	6 ABP97376	Human kie
19	100.5	11.8	1207	6 ABP97377	Human kie
20	100.5	11.8	1251	6 ABP97375	Human kie
21	100.5	11.8	1342	6 ABP97379	Human kie
22	100.5	11.8	1477	6 ABP97371	Human kie
23	100.5	11.8	1512	6 ABP97372	Human kie
24	100.5	11.8	1535	6 ABP97374	Human kie
25	100.5	11.8	1570	6 ABP97373	Human kie

26	100.5	11.8	1593	6 ABP97369	Human kie
27	100.5	11.8	1628	6 ABP97368	Human kie
28	100	11.7	3046	8 ADP31684	Human sec
29	97	11.3	434	4 ABB66756	Drosophil
30	96	11.2	623	5 ABB77796	Amino aci
31	96	11.2	1050	5 ABB77795	Antigen e
32	96	11.2	2108	8 ADN96130	Human NOV
33	95.5	11.2	193	2 AAY08491	Rat serin
34	95.5	11.2	216	2 AAY08492	Rat serin
35	95.5	11.2	405	5 ABB77797	Amino aci
36	95.5	11.2	1359	8 ADP31596	Human sec
37	95	11.1	585	8 ADP31445	Human sec
38	95	11.1	917	8 ADP30890	Human sec
39	94	11.0	148	2 AAY09280	YL-4 ther
40	93.5	10.9	1812	8 ADP31556	Human sec
41	93.5	10.9	1911	8 ADP31555	Human sec
42	92.5	10.8	2321	2 AAW49698	Human Not
43	92.5	10.8	2321	7 ADJ69828	Human hea
44	92.5	10.8	2321	8 ADH34621	Notch hom
45	92.5	10.8	2321	8 ADJ75570	Marker ge

ALIGNMENTS

RESULT 1

ADP02855

ID ADP02855 standard; protein; 148 AA.

XX AC ADP02855;

XX XX 09-SEP-2004 (first entry)

XX DE Solanum americanum proteinase inhibitor II protein SapIN2a.

XX KW proteinase inhibitor II; transformed plant; resistance; insect; pest;
 XX KW pathogen; programmed cell death; senescence.

XX OS Solanum americanum.

XX PN WO2004050873-A1.

XX PD 17-JUN-2004.

XX PF 01-DEC-2003; 2003WO-CN001020.

XX PR 29-NOV-2002; 2002US-0429992P.

XX PA (UYHK-) UNIV HONG KONG.

XX PI Chye M, Xu Z, Sin S;

XX DR WPI: 2004-450731/42.

XX DR N-PSDB; ADP02854.

XX PT New proteinase inhibitor, SapIN2a or SapIN2b, useful in producing
 XX PT transformed plants having enhanced resistance to insects, pest or
 XX PT pathogens and in which programmed cell death or senescence is inhibited.

XX PS Claim 5; SEQ ID NO 2; 90pp; English.

XX CC The invention relates to an isolated proteinase inhibitor II nucleic acid
 XX CC molecule (I). The proteinase inhibitor II nucleic acid molecule and the
 XX CC encoded polypeptide, methods are useful in producing transformed plants
 XX CC having enhanced resistance to insects, pest or pathogens and in which
 XX CC programmed cell death or senescence is inhibited. This sequence
 XX CC corresponds to the American black nightshade SapIN2a protein.

XX SQ Sequence 148 AA;

Query Match 100.0%; Score 855; DB 8; Length 148;
 Best Local Similarity 100.0%; Pred. No. 3e-64;
 Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MAVHKVSFLACLIVLGMFLAKHVDAKACTRECGHFSYGCIPRSGSPQKPICTNCCSG 60
Db |||||
QY 61 YKGCNYSAGDILICGESDPRNPKDCTFECDTQIAYSKCPSEGKMIKPTGCTTCCTG 120
Db |||||
QY 121 YQCYFYFDQDGFVCGESPEPKTTAYF 148
Db |||||
  RESULT 2
ADP02857
ID ADP02857 standard; protein; 152 AA.
AC ADP02857;
XX
DT 09-SEP-2004 (first entry)
XX
DE Solanum americanum proteinase inhibitor II protein SapIN2b.
XX
KW proteinase inhibitor II; transformed plant; resistance; insect; pest;
KW pathogen; programmed cell death; senescence.
XX
OS Solanum americanum.
XX
PN WO2004050873-A1.
XX
PD 17-JUN-2004.
XX
PF 01-DEC-2003; 2003WO-CN001020.
XX
PR 29-NOV-2002; 2002US-0429992P.
XX
PA (UYHK-) UNIV HONG KONG.
XX
PI Chye M, Xu Z, Sin S;
XX
DR WPI; 2004-450731/42.
DR N-PSDB; ADP02856.
XX
PT New proteinase inhibitor, SapIN2a or SapIN2b, useful in producing
PT transformed plants having enhanced resistance to insects, pest or
PT pathogens and in which programmed cell death or senescence is inhibited.
XX
PS Claim 6; SEQ ID NO 4; 90pp; English.
XX
CC The invention relates to an isolated proteinase inhibitor II nucleic acid
CC molecule (I). The proteinase inhibitor II nucleic acid molecule and the
CC encoded polypeptide, methods are useful in producing transformed plants
CC having enhanced resistance to insects, pest or pathogens and in which
CC programmed cell death or senescence is inhibited. This sequence
CC corresponds to the American black nightshade SapIN2b protein.
XX
SQ Sequence 152 AA;
  Query Match 71.9%; Score 615; DB 8; Length 152;
  Best Local Similarity 75.3%; Pred. No. 4.7e-44;
  Matches 110; Conservative 8; Mismatches 24; Indels 4; Gaps 2;
  QY 1 MAVHK-VSFLACLIVLGMFL- LAKHVDAKACTRECGHFSYGCIPRSGSPQKPICTN 56
  Db |||||
  QY 1 MAYHKEVSSLAYLLVLGLMFLHVS AVKHVDKPCPTRECGNLGYGICPRSGSPENP ICTN 60
  Db |||||
  QY 57 CCSGYKGCNYSAGDILICGESDPRNPKDCTFECDTQIAYSKCPSEGKMIKPTGCTT 116
  Db |||||
  QY 61 CCSGYKGCNYSANGTFCIGSSDPRNPKDCTFECDTQIAYSKCPSEGKMIKPTGCTT 120
  QY 117 CTTGYQCYFYFDQDGFVCGESPEP 142
  |||||

```

```

Db 121 CTTGYKGCYFYFSKSGEFVCEGSDEP 146
  RESULT 3
AAR20099
ID AAR20099 standard; protein; 153 AA.
XX
AC AAR20099;
XX
DT 24-MAR-1992 (first entry)
XX
DE Wound-inducible potato protease inhibitor IIK.
XX
KW methyl jasmonate; transgenic plant; predator defence protein.
XX
OS Solanum tuberosum.
XX
FH Key Location/Qualifiers
FT Peptide 1..30
FT /label= transit
FT Protein 31..153
FT /label= Inhibitor_IIK
XX
PN WO9118512-A.
XX
PD 12-DEC-1991.
XX
PF 25-MAY-1990; 90US-00528956.
XX
PR 25-MAY-1990; 90US-00528956.
XX
PA (UNIW ) WASHINGTON STATE UN.
XX
PI Ryan CA, Farmer EE;
XX
DR WPI; 1992-007120/01.
DR N-PSDB; AAQ20253.
XX
PT Induction of defence protein prodn. in plants - by treatment with agent
PT capable of such induction, esp. Jasmonic acid cpd.
XX
PS Example 7; Fig 6; 40pp; English.
XX
CC A Russett Burbank potato genomic library was screened by using nick-
CC translated wound-induced tomato inhibitor II cDNA as a probe. Positive
CC clones were rescreened and a clone contg. an 8kbp EcoRI insert hybridised
CC most strongly with the probe. A TaqI 2.6kbp fragment of this insert was
CC subcloned in M13 and a 1.24kbp fragment sequenced. The amino acid
CC sequence of the inhibitor was deduced from the nucleotide sequence. The
CC nascent protein is assumed to be processed during or after synthesis
CC between amino acid residues 30 and 31 to produce the mature inhibitor
XX
SQ Sequence 153 AA;
  Query Match 68.5%; Score 585.5; DB 2; Length 153;
  Best Local Similarity 70.0%; Pred. No. 1.4e-41;
  Matches 105; Conservative 14; Mismatches 28; Indels 3; Gaps 3;
  QY 1 MAVHK-VSFLACLIVLGMFL- AKHVDKACTRECGHFSYGCIPRSGSPQKPICTN 57
  Db |||||
  QY 1 MDVHKVEVFAVLLIIVLGLLVLSAMDVDAKACIRECGNLGFGICPRSGSPENP ICTN 60
  Db |||||
  QY 58 CCSGYKGCNYSAGDILICGESDPRNPKDCTFECDTQIAYSKCPSEGKMIKPTGCTTC 117
  Db |||||
  QY 61 CAGYKGCNYSANGATFCIGQSDPKPKACPLNCDPHIAYSKCPSEGKSLIYPTGCTTC 120
  QY 118 CTTGYQCYFYFDQDGFVCGESPEPKTTAY 147
  Db |||||
  Db 121 CTTGYKGCYFYFGKNGKFCVCGESDEPKANMY 150
  RESULT 4
AAR54135

```

ID AARS4135 standard; protein; 368 AA.
 AC AARS4135;
 XX
 DT 25-MAR-2003 (revised)
 DT 02-MAR-1995 (first entry)
 XX
 DE Deduced sequence of mature N-alata PI precursor.
 XX
 KW Type II serine proteinase inhibitor precursor; PI; tobacco;
 KW transgenic plant; anti-pathogen; anti-predator.
 KW Nicotiana alata.
 OS
 XX
 FH Key Location/Qualifiers
 FT Domain 1..58
 FT /label= 1
 FT Peptide 1..24
 FT /label= 1
 FT Active-site 5..6
 FT /label= reactive site
 FT Peptide 25..82
 FT /label= 2
 FT Domain 59..116
 FT /label= 2
 FT Active-site 63..64
 FT /label= reactive site
 FT Peptide 83..140
 FT /label= 3
 FT Domain 117..174
 FT /label= 3
 FT Active-site 121..122
 FT /label= reactive site
 FT Peptide 141..198
 FT /label= 4
 FT Domain 175..232
 FT /label= 4
 FT Active-site 179..180
 FT /label= reactive site
 FT Peptide 199..256
 FT /label= 5
 FT Domain 233..290
 FT /label= 5
 FT Active-site 237..238
 FT /label= reactive site
 FT Peptide 257..314
 FT /label= 6
 FT Domain 291..343
 FT /label= 6
 FT Active-site 295..296
 FT /label= reactive site
 FT Peptide 315..368
 FT /label= 7
 XX
 PN W09413810-A1.
 XX
 PD 23-JUN-1994.
 XX
 PF 16-DEC-1993; 93WO-AU000659.
 XX
 PR 16-DEC-1992; 92AU-00006399.
 XX
 XX (UYME) UNIV MELBOURNE.
 XX
 PI Anderson MA, Atkinson AH, Heath RL, Clarke AE;
 XX
 DR WPI; 1994-217886/26.
 DR N-PSDB; AAQ68728, AAQ68729.
 XX
 PT Nicotiana alata type II serine protease inhibitor precursor and DNA -
 PT useful in prodn of anti-pathogen or anti-predator constructs for plants.
 XX
 PS Claim 16; Page 45-47; 83pp; English.

XX A cDNA library, prepd. from mRNA from the stigmas and styles of mature
 CC flowers of N. alata was screened for clones of highly expressed genes
 CC which were not associated with self-incompatibility genotype. Clones
 CC encoding a protein with some identity to the type II proteinase
 CC inhibitors from potato and tomato were selected. The largest clone, NA-PI
 CC -2, is given in AAQ68729. The predicted AA sequence in AARS4135, AAQ68728
 CC is the coding region of AAQ68729. The type II serine PI has six conserved
 CC domains wherein the first two are 100% identical and contain chymotrypsin
 CC -specific sites. The 3rd, 4th and 5th domains share 95-98% identity and
 CC have sites specific for trypsin. The 6th domain also has a trypsin
 CC specific site but less identity to the 3rd, 4th and 5th domains (79-90%)
 CC due mainly to a divergent 3' sequence. It has a mol. wt. of approx. 42-
 CC 45kDa with an approx. 29 AA signal sequence. The N-terminal sequence of
 CC the monomeric PI is represented in each of the six repeated domains in
 CC the predicted sequence of the PI precursor protein. Thus, it is likely
 CC that the PI precursor protein is cleaved at six sites to produce seven
 CC peptides. Six of the seven peptides, peptides 2-7 (AARS4125-R54130) would
 CC be in the same mol. wt. range as the monomeric PI (about 6kDa) and would
 CC have the same N-terminal sequence. Peptide 7 does not contain a consensus
 CC site for trypsin or chymotrypsin. Peptide 1 (AARS4124) is small than
 CC 6kDa, has a different N-terminus and was not detected in a purified
 CC monomeric PI prepn. It could be envisaged that peptide 1 and peptide 7
 CC would form a functional proenzyme inhibitor with the inhibitory site on
 CC peptide 1 held in the correct conformation by disulphide bonds between
 CC the two peptides. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 368 AA;
 Query Match 44.3%; Score 379; DB 2; Length 368;
 Best Local Similarity 58.1%; Pred. No. 7.7e-24;
 Matches 68; Conservative 9; Mismatches 36; Indels 4; Gaps 3;
 QY 28 KACTREG-HFSYICPRSEGSFQKPICTNCCSGYKGNYYSAKDLICGESDPRNPKD 86
 DB 1 KACTLNCDPRIAYGVCPRSEKKNDRICTNCCAGTKGCKYFSDDTFVCEGESDPRNPKA 60
 QY 87 CTPECDTQIAYSKCPRSEKGMIIKPTGCTTCCTGYOGCYFDDGDGDFVCEGESPEPK 143
 DB 61 CTLNCDPRIAYGVCPRESEKK--NDRICTNCCAGTKGCKYFSDDTFVCEGES--DPR 114
 RESULT 5
 ADI56864
 ID ADI56864 standard; protein; 368 AA.
 XX
 AC ADI56864;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Plant defensin NaPI mature domain.
 XX
 KW insecticide; antifungal; virucide; antibacterial; flower development;
 KW plant pest resistance; floral defensin precursor; transgenic plant;
 KW insect resistance.
 XX
 OS Unidentified.
 XX
 PN W0200263011-A1.
 XX
 PD 15-AUG-2002.
 XX
 PF 08-FEB-2002; 2002WO-AU000123.
 XX
 PR 08-FEB-2001; 2001US-0267271P.
 XX
 PA (HEXI-) HEXIMA LTD.
 XX
 PI Anderson MA, Lay FT, Heath RL;
 XX
 DR WPI; 2002-657538/70.
 DR N-PSDB; ADI56863.
 XX

PT New isolated plant floral defensin-like polypeptides and polynucleotides,
 PT for generating transgenic plants having resistance or at least reduced
 PT sensitivity to plant pests including insects, microorganisms, fungi
 and/or viruses.

XX Disclosure; SEQ ID NO 57; 164pp; English.

XX The invention relates to an isolated polypeptide comprising, in its
 CC precursor form, an N-terminal signal domain, a mature domain and an
 CC acidic C-terminal domain, where the polypeptide is produced during flower
 CC development and its mature domain has activity against one or more plant
 CC pests. Plant floral defensin-like polypeptides and polynucleotides are
 CC useful in generating transgenic plants having resistance or at least
 CC reduced sensitivity to plant pests including insects, microorganisms,
 CC fungi and/or viruses. They are also useful in generating recombinant
 CC defensin-like molecules for use in the topical application of
 CC compositions to prevent or retard pest-infestation of plants. The floral-
 CC and seed-derived defensins are useful in the generation of insect
 CC resistance in plants. This sequence represents the amino acid sequence of
 CC the mature domain from a plant defensin protein.

XX Sequence 368 AA;

Query Match 44.3%; Score 379; DB 5; Length 368;
 Best Local Similarity 58.1%; Pred. No. 7,7e-24;
 Matches 68; Conservative 9; Mismatches 36; Indels 4; Gaps 3;

QY 28 KACTREG-HFSYIGIPRSGSPQKPICTNCCSGYKGNYSAGDLCIGESDPRNPKD 86
 Db 1 KACTLNCDPRIAYGVCPRSEKKNDRICTNCCAGTCKRYFDDGTFVCEGESDPRNPKA 60
 QY 87 CTPECDTQIAYSKCPRSEGMIIKPTGCTTCCTGYOGCYFFDQDGFVCEGESPEPK 143
 Db 61 CTLNCDPRIAYGVCPRSEKK--NDRICTNCCAGTCKRYFDDGTFVCEGES-DPR 114

RESULT 6

AAR54125
 ID AAR54125 standard; peptide; 58 AA.

XX AAR54125;

XX 25-MAR-2003 (revised)

DT 02-MAR-1995 (first entry)

XX Sequence of peptide 2 of N-alata PI precursor.

XX Type II serine proteinase inhibitor precursor; PI; tobacco;
 KW transgenic plant; anti-pathogen; anti-predator; peptide.

XX Nicotiana alata.

XX WO9413810-A1.

XX 23-JUN-1994.

XX 16-DEC-1993; 93WO-AU000659.

XX 16-DEC-1992; 92AU-00006399.

XX (UYME) UNIV MELBOURNE.

XX Anderson MA, Atkinson AH, Heath RL, Clarke AE;

XX WPI; 1994-217886/26.

DR N-PSDB; AAQ68728, AAQ68729.

XX Nicotiana alata type II serine protease inhibitor precursor and DNA -
 PT useful in prodn of anti-pathogen or anti-predator constructs for plants.

XX Claim 18; Page 50; 83pp; English.

XX A cDNA library, prepd. from mRNA from the stigmas and styles of mature

CC flowers of N. alata was screened for clones of highly expressed genes
 CC which were not associated with self-incompatibility genotype. Clones
 CC encoding a protein with some identity to the type II proteinase
 CC inhibitors from potato and tomato were selected. The largest clone, NA-PI
 CC -2, is given in AAQ68729. The predicted AA sequence in AAR54135, AAQ68728
 CC domains wherein the first two are 100% identical and contain chymotrypsin
 CC -specific sites. The 3rd, 4th and 5th domains share 95-98% identity and
 CC have sites specific for trypsin. The 6th domain also has a trypsin
 CC specific site but less identity to the 3rd, 4th and 5th domains (79-90%)
 CC due mainly to a divergent 3' sequence. It has a mol. st. of approx. 42-
 CC 45kDa with an approx. 29 AA signal sequence. The N-terminal sequence of
 CC the monomeric PI is represented in each of the six repeated domains in
 CC the predicted sequence of the PI precursor protein. Thus, it is likely
 CC that the PI precursor protein is cleaved at six sites to produce seven
 CC peptides. Six of the seven peptides, peptides 2-7 (AAR54125-R54130) would
 CC be in the same mol. wt. range as the monomeric PI (about 6kDa) and would
 CC have the same N-terminal sequence. Peptide 7 does not contain a consensus
 CC site for trypsin or chymotrypsin. Peptide 1 (AAR54124) is small than
 CC 6kDa, has a different N-terminus and was not detected in a purified
 CC monomeric PI prepn. It could be envisaged that peptide 1 and peptide 7
 CC would form a functional proenzyme inhibitor with the inhibitory site on
 CC peptide 1 held in the correct conformation by disulphide bonds between
 CC the two peptides. A monomer which corresp. to any of the peptides 2-7 is
 CC claimed. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 58 AA;

Query Match 26.2%; Score 224; DB 2; Length 58;
 Best Local Similarity 68.5%; Pred. No. 1.4e-11;
 Matches 37; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 53 ICTNCCSGYKGNYSAGDLCIGESDPRNPKDCTFECDTQIAYSKCPRSEK 106
 Db 3 ICTNCCAGTCKRYFDDGTFVCEGESDPRNPKACTLNCDPRIAYGVCPRSEK 56

RESULT 7

AAR54129

ID AAR54129 standard; peptide; 58 AA.

XX AAR54129;

XX 25-MAR-2003 (revised)

DT 02-MAR-1995 (first entry)

XX Sequence of peptide 6 of N-alata PI precursor.

XX Type II serine proteinase inhibitor precursor; PI; tobacco;
 KW transgenic plant; anti-pathogen; anti-predator; peptide.

XX Nicotiana alata.

XX WO9413810-A1.

XX 23-JUN-1994.

XX 16-DEC-1993; 93WO-AU000659.

XX 16-DEC-1992; 92AU-00006399.

XX (UYME) UNIV MELBOURNE.

XX Anderson MA, Atkinson AH, Heath RL, Clarke AE;

XX WPI; 1994-217886/26.

DR N-PSDB; AAQ68728, AAQ68729.

XX Nicotiana alata type II serine protease inhibitor precursor and DNA -
 PT useful in prodn of anti-pathogen or anti-predator constructs for plants.

XX Claim 18; Page 52; 83pp; English.

XX

RESULT 15
ADP30830
ID ADP30830 standard; protein; 186 AA.
XX AC
XX ADP30830;
XX DT
XX 12-AUG-2004 (first entry)
XX DE Human secreted protein SEQ ID #1597.
XX DE Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide;
KW cancer; inflammatory; immune; human secreted protein.
XX KW
XX Homo sapiens.
OS
XX WO2004035732-A2.
XX PN
XX PD
XX 29-APR-2004.
XX PF
XX 28-AUG-2003; 2003WO-US025780.
XX PR
XX 29-AUG-2002; 2002US-0406576P.
XX PR
XX 29-AUG-2002; 2002US-0406579P.
XX PR
XX 29-AUG-2002; 2002US-0406585P.
XX PR
XX 29-AUG-2002; 2002US-0406588P.
XX PR
XX 29-AUG-2002; 2002US-0406608P.
XX PR
XX 29-AUG-2002; 2002US-0406611P.
XX PR
XX 29-AUG-2002; 2002US-0406612P.
XX PR
XX 29-AUG-2002; 2002US-0406616P.
XX PR
XX 29-AUG-2002; 2002US-0406640P.
XX PR
XX 29-AUG-2002; 2002US-0406642P.
XX PR
XX 29-AUG-2002; 2002US-0406646P.
XX PR
XX 29-AUG-2002; 2002US-0406653P.
XX PR
XX 29-AUG-2002; 2002US-0406655P.
XX PR
XX 29-AUG-2002; 2002US-0406666P.
XX PR
XX 17-SEP-2002; 2002US-0410946P.
XX PR
XX 17-SEP-2002; 2002US-0410947P.
XX PR
XX 17-SEP-2002; 2002US-0410948P.
XX PR
XX 17-SEP-2002; 2002US-0410949P.
XX PR
XX 17-SEP-2002; 2002US-0410953P.
XX PR
XX 17-SEP-2002; 2002US-0410957P.
XX PR
XX 17-SEP-2002; 2002US-0410958P.
XX PR
XX 17-SEP-2002; 2002US-0410959P.
XX PR
XX 17-SEP-2002; 2002US-0410960P.
XX PR
XX 17-SEP-2002; 2002US-0410961P.
XX PR
XX 17-SEP-2002; 2002US-0410962P.
XX PR
XX 17-SEP-2002; 2002US-0411019P.
XX PR
XX 17-SEP-2002; 2002US-0411022P.
XX PR
XX 17-SEP-2002; 2002US-0411023P.
XX PR
XX 17-SEP-2002; 2002US-0411024P.
XX PR
XX 17-SEP-2002; 2002US-0411032P.
XX PR
XX 17-SEP-2002; 2002US-0411035P.
XX PR
XX 17-SEP-2002; 2002US-0411037P.
XX PR
XX 17-SEP-2002; 2002US-0411041P.
XX PR
XX 17-SEP-2002; 2002US-0411045P.
XX PR
XX 17-SEP-2002; 2002US-0411046P.
XX PR
XX 17-SEP-2002; 2002US-0411048P.
XX PR
XX 17-SEP-2002; 2002US-0411052P.
XX PR
XX 17-SEP-2002; 2002US-0411055P.
XX PR
XX 17-SEP-2002; 2002US-0411073P.
XX PR
XX 17-SEP-2002; 2002US-0411082P.
XX PR
XX 17-SEP-2002; 2002US-0411101P.
XX PR
XX 17-SEP-2002; 2002US-0411111P.
XX PR
XX 18-APR-2003; 2003US-0463700P.
XX PR
XX 18-APR-2003; 2003US-0463708P.
XX PR
XX 18-APR-2003; 2003US-0463716P.
XX PR
XX 18-APR-2003; 2003US-0463732P.
XX PR
XX 02-MAY-2003; 2003US-0467199P.
XX PR
XX 02-MAY-2003; 2003US-0467201P.
XX PR
XX 02-MAY-2003; 2003US-0467203P.
XX PR
XX 02-MAY-2003; 2003US-0467230P.
XX PR
XX 19-MAY-2003; 2003US-0471306P.
XX PR
XX 19-MAY-2003; 2003US-0471336P.

PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
XX
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Huang MM, Kothakota S, Haislan L, Linnemann T;
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
XX
XX WPI; 2004-348438/32.
XX
XX New nucleic acid molecule for diagnosing, preventing or treating diseases
PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
PT genetic, bacterial and viral diseases.
XX
XX Claim 1; SEQ ID NO 2828; 428pp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a polypeptide which is believed to be cytosolic,
CC antiinflammatory, immunosuppressive, antibacterial and virucidal. The
CC composition and methods are useful for diagnosing, preventing and
CC treating diseases such as proliferative (e.g. cancer), inflammatory,
CC immune, metabolic, genetic, bacterial and viral diseases. The present
CC sequence represents a human secreted protein. The present sequence is
CC available on WIPOWEB and is not in the specification.
XX
XX Sequence 186 AA;
Query Match 12.2%; Score 104.5; DB 8; Length 186;
Best Local Similarity 31.4%; Pred. No. 0.48;
Matches 32; Conservative 4; Mismatches 43; Indels 23; Gaps 5;
QY 29 ACTREC-GHFSYGICPRSGSPQKICTNCCSGYKGCNYYSAKG-----DLICEGESDPR 82
Db 70 ACTACAGGGGGGCGAGGAGA----TCTTCGACGGCGCATGGACCCCTATCTGTCCAC 125
QY 83 NPKDCTFECDTQIAYSKCPRSEKMIKPTGTCTTCTCTGYGC 124
Db 126 ATGGCCTGCCTCTCCTTC-----TTCTTCCTG--GC 154
Search completed: January 24, 2005, 16:48:48
Job time : 69 secs

THIS PAGE IS BLANK